12 September 2008 | $10

Science

AAAS

COVER

Millions of books written before the computer era are being digitized for preservation. Because the ink has faded, optical character recognition software cannot decipher many words. Through a repurposing of an existing online security technology called CAPTCHA, these words are being manually transcribed by millions of Web users. See page 1465.

Photo: Joshua Franzos

DEPARTMENTS

1411 *Science* Online
1413 This Week in *Science*
1419 Editors' Choice
1422 Contact *Science*
1425 Random Samples
1427 Newsmakers
1511 New Products
1513 *Science* Careers

EDITORIAL

1417 **Mediterranean Scientopolitics**
by Ahmed Zewail

NEWS OF THE WEEK

U.K. Education Reform: Too Much of a Good Thing? 1428
Brainy Babies and Risky Births for Neandertals 1429
India Hopes New Fellowships Will Attract Expat Scientists 1431
SCIENCESCOPE 1431
Broad Gives $400 Million More to Cambridge Institute 1432
McCain, Obama Present Their Wars on Cancer 1432
Quantum Flashlight Pierces the Darkness With a Few Percent as Many Photons 1433
>> *Report p. 1463*

NEWS FOCUS

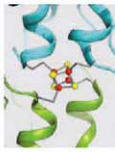
The Mushroom Cloud's Silver Lining 1434
Forgers Face the Nuclear Option
>> *Science Podcast*
Obama and McCain Are Swept Up in a Surprising Space Race 1438
New Institute Shoots for the Moon
The Houbara: Headed for Oblivion? 1441

LETTERS

Working the Crowd *A. Gaggioli and G. Riva* 1443
Southern Ocean Not So Pristine
L. K. Blight and D. G. Ainley
Diminishing Sea Ice *G. C. Ray* et al.
Response *B. S. Halpern* et al.
Microscopy for Life Scientists *N. J. Fullwood*
Archaeology Without Borders *S. K. Basu*
CORRECTIONS AND CLARIFICATIONS 1446

BOOKS *ET AL.*

Worlds Before Adam The Reconstruction of Geohistory in the Age of Reform 1447
M. J. S. Rudwick, reviewed by R. J. O'Connor
The Animal Research War 1448
P. M. Conn and J. V. Parker, reviewed by D. C. Runkle


1452

POLICY FORUM

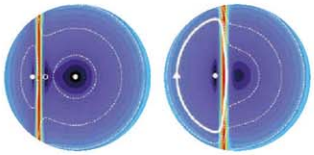
Do We Need "Synthetic Bioethics"? 1449
E. Parens, J. Johnston, J. Moses

PERSPECTIVES

Return to the Proliferative Pool 1450
A. González-Reyes and J. Casanova >> *Report p. 1496*
Dynamics of Body Size Evolution 1451
K. Roy
Bringing Stability to Highly Reduced Iron–Sulfur Clusters 1452
E. Münck and E. L. Bominaar
Understanding Soil Time 1454
S. L. Brantley >> *Perspective p. 1455*
An Uncertain Future for Soil Carbon 1455
S. E. Trumbore and C. I. Czimczik >> *Perspective p. 1454*

CONTENTS continued >>

Science



SCIENCE EXPRESS

www.sciencexpress.org

GENETICS

Species-Specific Transcription in Mice Carrying Human Chromosome 21

M. D. Wilson et al.

An aneuploid mouse carrying a human chromosome shows that genetic sequence can dominate epigenetic, cellular, and organismal effects in determining transcriptional regulation and gene expression.

10.1126/science.1160930

CLIMATE CHANGE

Atmospheric CO_2 and Climate on Millennial Time Scales During the Last Glacial Period

J. Ahn and E. J. Brook

A detailed gas record from the Byrd ice core from 90,000 to 20,000 years ago shows that warming episodes tracked high CO_2 levels in Antarctica but lagged by several thousands of years in Greenland.

10.1126/science.1160832

APPLIED PHYSICS

Cavity Optomechanics with a Bose-Einstein Condensate

F. Brennecke, S. Ritter, T. Donner, T. Esslinger

Coupling a Bose-Einstein condensate to an optical cavity holding a few trapped photons provides a sensitive probe of mechanical oscillations in the quantum regime.

10.1126/science.1163218

CLIMATE CHANGE

Northern Hemisphere Controls on Tropical Southeast African Climate During the Past 60,000 Years

J. E. Tierney et al.

Abrupt changes in precipitation and temperature resolved in a record spanning the past 60,000 years from Lake Tanganyika, East Africa, are coeval with Northern Hemisphere climate events.

10.1126/science.1160485

TECHNICAL COMMENT ABSTRACTS

ECOLOGY

Comment on "A Global Map of Human Impact on Marine Ecosystems" 1446

M. R. Heath

full text at www.sciencemag.org/cgi/content/full/321/5895/1446b

Response to Comment on "A Global Map of Human Impact on Marine Ecosystems"

K. A. Selkoe et al.

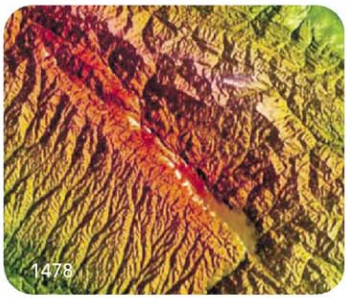
full text at www.sciencemag.org/cgi/content/full/321/5895/1446c

REVIEW

APPLIED PHYSICS

Cooling, Heating, Generating Power, and Recovering Waste Heat with Thermoelectric Systems 1457

L. E. Bell



BREVIA

GENETICS

A Mutation in Hairless Dogs Implicates *FOXI3* in Ectodermal Development 1462

C. Drögemüller et al.

Mutations in a transcription factor gene involved in ectodermal development cause a lack of hair and abnormal teeth in Chinese Crested, Mexican, and Peruvian hairless dogs.

REPORTS

PHYSICS

Enhanced Sensitivity of Photodetection via Quantum Illumination 1463

S. Lloyd

Quantum-mechanically entangled light, in which one photon is kept as a reference, can exponentially improve the imaging of an object, as compared with unentangled illumination. >> *News story p. 1433*

COMPUTER SCIENCE

reCAPTCHA: Human-Based Character Recognition via Web Security Measures 1465

L. von Ahn et al.

A security system that relies on the superior performance of humans in comparison to computers in reading distorted text can be harnessed for digitized scanned documents.

MATERIALS SCIENCE

A Rubberlike Stretchable Active Matrix Using Elastic Conductors 1468

T. Sekitani et al.

A carbon nanotube–polymer film containing organic transistors and coated with silicon rubber can maintain its electrical properties while being stretched up to 70 percent.

CONTENTS continued >>

REPORTS *CONTINUED...*

MATERIALS SCIENCE

Imaging of Transient Structures Using Nanosecond in Situ TEM 1472
J. S. Kim et al.
Rapidly pulsing electrons through a transmission electron microscope allows imaging of localized cooling and phase separation along a propagating reaction front in a laminate. >> Science *Podcast*

PLANETARY SCIENCE

The Magnetic Memory of Titan's Ionized Atmosphere 1475
C. Bertucci et al.
Cassini observations show that Saturn's moon Titan retains an imprinted memory of Saturn's magnetic field above its ionosphere, even after passing outside the field.

GEOPHYSICS

Postseismic Relaxation Along the San Andreas Fault at Parkfield from Continuous Seismological Observations 1478
F. Brenguier et al.
Correlating 5 years of seismic noise among nearby receivers reveals subtle seismic velocity signals reflecting changes in the properties of the San Andreas fault at Parkfield.

ATMOSPHERIC SCIENCE

Atmospheric Warming and the Amplification of Precipitation Extremes 1481
R. P. Allan and B. J. Soden
Satellite data show that in the tropics, heavy rain events have increased in warmer months and decreased in colder months, more than predicted by climate models.

PALEONTOLOGY

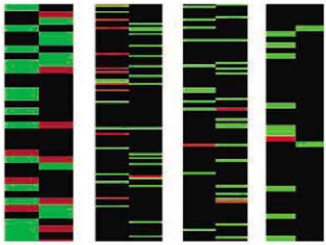
Superiority, Competition, and Opportunism in the Evolutionary Radiation of Dinosaurs 1485
S. L. Brusatte, M. J. Benton, M. Ruta, G. T. Lloyd
During their early radiation, dinosaur morphology evolved at comparable rates to that of competing archosaurs, implying that opportunity, not superiority, influences their success.

ECOLOGY

Niche Partitioning Increases Resource Exploitation by Diverse Communities 1488
D. L. Finke and W. E. Snyder
In an ecosystem comprising a parasite, an aphid, and a radish, the use of different resources by each species, not species diversity per se, increases overall consumption.

MOLECULAR BIOLOGY

Degradation of microRNAs by a Family of Exoribonucleases in *Arabidopsis* 1490
V. Ramachandran and X. Chen
A class of nucleases specific for short single-stranded RNAs is found to degrade microRNAs in *Arabidopsis*; their mutation results in numerous developmental defects.



1499

MEDICINE

Activation of Aldehyde Dehydrogenase-2 Reduces Ischemic Damage to the Heart 1493
C.-H. Chen et al.
A compound that activates the mitochondrial enzyme aldehyde dehydrogenase-2 reduces the extent of heart damage in a rodent model of heart attack. >> Science *Podcast*

DEVELOPMENTAL BIOLOGY

Dual Origin of Tissue-Specific Progenitor Cells in *Drosophila* Tracheal Remodeling 1496
M. Weaver and M. A. Krasnow
When fruit flies metamorphose from larvae, a new trachea forms both from undifferentiated cells of the imaginal disc and differentiated cells that re-enter the cell cycle.
>> *Perspective p. 1450*

CELL BIOLOGY

FBXW7 Targets mTOR for Degradation and Cooperates with PTEN in Tumor Suppression 1499
J.-H. Mao et al.
A tumor suppressor is shown to control the degradation of a central protein regulator of cell proliferation.

NEUROSCIENCE

Unsupervised Natural Experience Rapidly Alters Invariant Object Representation in Visual Cortex 1502
N. Li and J. J. DiCarlo
Neurons in the most complex area of the brain's visual cortex can respond to a particular object in any orientation by rapidly learning to associate multiple views of that object.

CELL BIOLOGY

Conformational Switch of Syntaxin-1 Controls Synaptic Vesicle Fusion 1507
S. H. Gerber et al.
The synaptic vesicle protein that mediates membrane fusion during exocytosis also regulates the rate and extent of this process by controlling vesicle tethering.

AAAS
ADVANCING SCIENCE, SERVING SOCIETY

SCIENCE (ISSN 0036-8075) **is published weekly on Friday, except the last week in December, by the American Association for the Advancement of Science, 1200 New York Avenue, NW, Washington, DC 20005.** Periodicals Mail postage (publication No. 484460) paid at Washington, DC, and additional mailing offices. Domestic individual membership and subscription (51 issues): $144 ($74 allocated to subscription). Domestic institutional subscription (51 issues): $970; Foreign postage extra: Mexico, Caribbean (surface mail) $55; other countries (air assist delivery) $85. First class, airmail, student, and emeritus rates on request. Canadian rates with GST available upon request, GST #125488322. Publications Mail Agreement Number 1069624. SCIENCE is printed on 30 percent post-consumer recycled paper. **Printed in the U.S.A.**

Change of address: Allow 4 weeks, giving old and new addresses and 8-digit account number. **Postmaster:** Send change of address to AAAS, P.O. Box 96178, Washington, DC 20090–6178. **Single-copy sales:** $10.00 current issue, $15.00 back issue prepaid includes surface postage; bulk rates on request. **Authorization to photocopy** material for internal or personal use under circumstances not falling within the fair use provisions of the Copyright Act is granted by AAAS to libraries and other users registered with the Copyright Clearance Center (CCC) Transactional Reporting Service, provided that $20.00 per article is paid directly to CCC, 222 Rosewood Drive, Danvers, MA 01923. The identification code for Science is 0036-8075. Science is indexed in the *Reader's Guide to Periodical Literature* and in several specialized indexes.

Printed on 30% post-consumer recycled paper.

CONTENTS continued >>

Homely but hardy.

SCIENCENOW

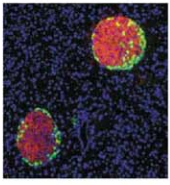
www.sciencenow.org

HIGHLIGHTS FROM OUR DAILY NEWS COVERAGE

"Water Bears" Survive Earth Orbit
Hardy little creatures shrug off extremes of life outside a spacecraft.

Finding a Guy Who Looks Like Dear Old Dad
Men and women choose partners who resemble their own parents.

A Knack for Numbers
Kids who are good at judging relative quantities are also good at math.



Pancreatic islets in p110β mutant mice.

SCIENCE SIGNALING

www.sciencesignaling.org

THE SIGNAL TRANSDUCTION KNOWLEDGE ENVIRONMENT

RESEARCH ARTICLE: Phosphoinositide 3-Kinase p110β Activity—Key Role in Metabolism and Mammary Gland Cancer but Not Development
E. Ciraolo, M. Iezzi, R. Marone, S. Marengo, C. Curcio, C. Costa, O. Azzolino, C. Gonella, C. Rubinetto, H. Wu, W. Dastrù, E. L. Martin, L. Silengo, F. Altruda, E. Turco, L. Lanzetti, P. Musiani, T. Rückle, C. Rommel, J. M. Backer, G. Forni, M. P. Wymann, E. Hirsch
The phosphoinositide 3-kinase p110β subunit has noncatalytic functions; its catalytic activity is pertinent to both diabetes and cancer.

PERSPECTIVE: Smad Signaling Dynamics—Insights from a Parsimonious Model
H. Shankaran and H. S. Wiley
Computational modeling of protein localization dynamics yields new information about Smad signaling.

GLOSSARY
Find out what Kir, Vg1, and YAP mean in the world of cell signaling.

How to deal with journalists.

SCIENCE CAREERS

www.sciencecareers.org/career_development

FREE CAREER RESOURCES FOR SCIENTISTS

Your Research in the Headlines—Dealing With the Media
E. Pain
Being prepared can improve the odds that your scientific work is portrayed accurately in the media.

Team Science and the Diversity Advantage
C. Rey
Do scientists with experiences in different cultures have an advantage in an era of team science?

In Person—Peter Brown, Patent Attorney
P. Brown
An award-winning New Zealand scientist tells why he decided to move from the bench to the bar.

International Grants and Fellowship Index
GrantsNet Staff
Learn about the latest funding opportunities from Europe, Asia, and the Americas.

SCIENCEPODCAST

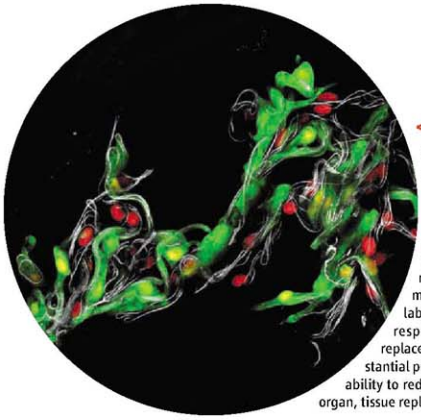
www.sciencemag.org/multimedia/podcast

FREE WEEKLY SHOW

Download the 12 September *Science* Podcast to hear about damage control for the heart, imaging transient phenomena, a silver lining of atomic bomb testing, and more.

CREDITS: (SCIENCENOW) GOLDSTEIN LAB/UNC CHAPEL HILL; (SCIENCE CAREERS) MARCIN WECHARY/CREATIVE COMMONS; (SCIENCE SIGNALING) ELISA CIRAOLO/UNIVERSITY OF TURIN/ITALY

Separate individual or institutional subscriptions to these products may be required for full-text access.



CREDITS (TOP TO BOTTOM): MOLLY WEAVER; BERTUCCI ET AL.

<< Tissue Remodeling

When organs are damaged, dying cells must be replaced to maintain organ function. Using insect metamorphosis as a model for tissue replacement, **Weaver and Krasnow** (p. 1496, published online 31 July; see the Perspective by **González-Reyes**) examine the progenitor cells that rebuild the fly respiratory system. For most fly organs, new tissue arises from undifferentiated progenitor cells associated with the organ that remain quiescent until metamorphosis when they proliferate, migrate over, and replace dying cells. Individual cells were labeled and their fates followed to identify a second population of respiratory progenitors that arise from differentiated cells and replace local regions of the airways. These differentiated cells have substantial proliferative potential and developmental plasticity, including the ability to redifferentiate as a new cell type. Thus, even in a simple epithelial organ, tissue replacement requires different progenitors and cellular strategies.

Recovering Wasted Heat

The recovery of waste heat, especially on the industrial scale, normally relies on transferring heat with a working fluid, such as converting liquid water to steam. The efficiency of these approaches is highest on a large scale and when thermal gradients are very high. For recovery of heat on a smaller scale, such as from a car engine, thermoelectric systems, which use electrical current as the working fluid, are more attractive. Although most of the attention in thermoelectric materials have focused on their intrinsic efficiency (described by the parameter ZT), **Bell** (p. 1457) reviews the engineering challenges and opportunities in using such materials in cars, electronics, and other applications for heating, cooling, and power generation.

Extreme Behavior

Global warming is expected to have a large effect on the amount and distribution of precipitation, with wet areas projected to become wetter and dry areas drier, and an overall increase in total rainfall. Another important aspect of these predicted changes is the frequency of extreme rainfall events, because the impact of a few heavy rain events is very different from that of many more moderate ones. **Allan and Soden** (p. 1481, published online 7 August) use satellite observations and model simulations to evaluate how climate warming is affecting the frequency and strength of rain events. Heavy rains are occurring with increasing frequency when it is warm and less often when it is cold, and these extremes are happening more frequently than models have suggested they should. This implies that the impacts of precipitation changes due to global warming could be greater than have been assumed.

GOTCHA?

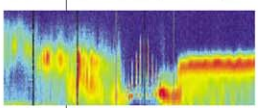
Thankfully, there are still tasks that humans can do that computers cannot. One of the 21st-century manifestations of this difference is the use of CAPTCHAs (distorted alphanumeric strings that must be read and typed) to safeguard entry into Web sites against nonhuman entities. **Von Ahn *et al.*** (p. 1465, published online 14 August; cover) describe a modification of this algorithm that serves to capture the effort expended by human users and to direct it toward digitizing scanned documents. Optical character recognition programs are unable to transcribe scans of printed matter for a variety of reasons, such as uneven shrinkage of the paper or fading of the ink; using these unrecognized words as queries for Web site visitors to decode exemplifies the approach known as crowd computing.

Bend Me, Stretch Me, Flex Me, Connect Me

One restriction in the development of rubbery electronics that can stretch as well as bend and flex is the need for a conductive, elastic material, especially for the interconnects between circuits. **Sekitani *et al.*** (p. 1468, published online 7 August) describe the development of a stretchable, multilayer single-walled carbon nanotube–polymer elastomer composite and incorporate it into an active matrix array of organic transistors. The elastomer shows excellent mechanical and electrical properties with both low resistivity and reversible stretchability to large levels of strain.

A Tale of Titan

Titan lacks its own internal magnetic field but is greatly affected by Saturn's magnetic field. Recently, Titan passed outside of the influence of Saturn's magnetic field, and was observed by the Cassini spacecraft, providing **Bertucci *et al.*** (p. 1475) a drastically different magnetic environment for comparison. Titan's ionosphere retained a "memory" of Saturn's field for some time, until it seems a magnetic reconnection in its tail replaced this field with the interplanetary one.



Dynamic Electron Microscopy

Transmission electron microscopy (TEM) is an excellent tool for studying structural changes in materials. While the resolution of the instrument has improved considerably, one challenge is being able to track fast-occurring phenomena with high resolution. Rapid snapshots can be acquired by using a laser to pump the electron gun. **Kim *et al.*** (p. 1472) apply this

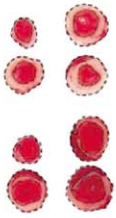
Continued on page 1415

Continued from page 1413

dynamic TEM method, where a second laser is used to initiate a reaction in a thin multilayer foil. Altering the time delay for the arrival of the electrons from the gun allowed for observation of localized cooling and phase separation at the propagating reaction front as the two materials mix and react with each other.

Rise of the Dinosaurs

Is the diversification of new groups of organism a matter of competitive superiority over other groups occupying similar niches, or does historical contingency play a part? **Brusatte *et al.*** (p. 1485) document the evolutionary patterns of the initial radiation of dinosaurs and other archosaur groups in the Triassic. For the first 30 million years of their history, dinosaurs lived alongside another major clade, the crurotarsan archosaurs, which occupied similar niches, exhibited a greater range of morphology, and evolved at indistinguishable rates. These findings cast doubt on long-standing ideas of dinosaur "superiority" and notions that dinosaurs were preordained for success from the start.



Damage Control for the Heart

Many forms of heart disease begin with an ischemic event during which the heart muscle receives an inadequate blood supply, resulting in the accumulation of toxic metabolites that cause irreversible tissue damage. Studying rodent models, **Chen *et al.*** (p. 1493) found that a mitochondrial enzyme called aldehyde dehydrogenase 2 (ALDH2) was consistently activated in hearts that were the most resistant to ischemia-induced damage. In a rat heart-attack model, administration of a small-molecule activator of ALDH2 (Alda-1) prior to the ischemic insult led to a reduction in the extent of heart damage, an effect most likely due to decreased formation of cytotoxic aldehydes. Thus, Alda-1 or related compounds potentially might be used therapeutically to minimize heart damage in controlled settings such as coronary bypass surgery.

You Are What You Eat?

Ecological models suggest that biodiversity arises from the partitioning of resources among species, allowing new species with unique resource-use patterns to invade communities. However, these models have not been tested empirically because real-world species differences in resource use are often confounded with other species traits (size, rate of growth, metabolic rate, etc.). **Finke and Snyder** (p. 1488) overcome these obstacles by exploiting host-fidelity behavior among a group of parasitoid wasps that attack aphids. While each wasp species is a generalist consumer that attacks many aphid species, individual wasps prefer to attack the same host species from which they themselves emerged. By rearing wasps of different species on each of several aphid species, consumer wasp communities were constructed that could be independently manipulated for consumer species identity, species richness, and patterns of resource use. Exploitation of the aphid resource clearly improved with greater consumer biodiversity, but only when constituent consumers were specialists with distinct resource-niche partitioning. Thus differences in resource use among species, rather than biodiversity per se, intensify resource exploitation at higher levels of consumer biodiversity.

Now You See It, Now You Don't

Each object can cast many different images on the eye. How can the brain combine different views of an object into a single object representation? Neurons at the inferior temporal cortex (brain area IT), the top processing level of the visual system, signal the presence of individual objects even if those objects appear in different positions. **Li and DiCarlo** (p. 1502) recorded neuronal responses in area IT of two monkeys to different objects presented at the central position and 3 degrees above or below. By systematically swapping object identity between two objects whenever the monkey made a fast eye movement (saccade) to one particular position in the visual field, the response of the IT neuron became less selective to the objects at the swap position or even inverted its selectivity. Thus, object representations in area IT can change in a short period of time.

CREDIT: CHEN *ET AL.*

Ahmed Zewail is the Linus Pauling Chair Professor at the California Institute of Technology and the 1999 Nobel Prize winner in Chemistry.

Mediterranean Scientopolitics

ON THIS YEAR'S BASTILLE DAY IN JULY, THE PRESIDENT OF FRANCE, NICOLAS SARKOZY, INAUgurated a new initiative for uniting the Mediterranean South with Europe in general, and France in particular. The aim of the Mediterranean Union (MU), an analogue of the post–Cold War European Union (EU), is to "lay the foundations of a political, economic and cultural union founded on the principles of strict equality." Comprising 27 EU members and states from the Middle East, North Africa, and the Balkans, the MU would in principle unite close to 800 million people. In June, a meeting was held at the Institut de France with representation from many academies, scientists, and politicians to discuss possible cooperative programs. The goals expressed at the meeting are admirable; however, the MU's motives need to be clearly defined, as the issues for the MU are very different from those for the EU. Most important, thus far missing in the fabric of the former is an explicit role for education and science.

The Mediterranean people have a rich history encompassing cradles of civilization ranging from Egypt and Greece to the Roman Empire. As the word implies in Latin, the Mediterranean was considered the "Middle Earth," but at present the disparity between North and South is alarming. The difference in gross domestic product between the two is staggering, and illiteracy, deterioration in education, and the unfavorable state of governance in the South have put many there at a disadvantage. Despite these challenges, the MU could redefine the state of North-South cooperation by providing new opportunities for progress—but only if differences and concerns are openly addressed.

The integration of Eastern and Western Europe is to some extent easier than that of the North and South Mediterranean because religions and cultures are more diverse in the Mediterranean Basin. For the initiative to succeed, the leaders in MU nations must promote economic and political strategies that respect these differences. The benefits of free trade and liberty, coupled with dialogues of cultures through scholarly discourse that promote mutual acceptance, will undoubtedly lead to stronger bonding among and security for the nations in the region. But if the MU is directed by political agendas, such as distancing Turkey from the EU or isolating particular countries in the EU or Arab League, it will ultimately unravel and become a medium for slogans and the polarization of nations. The main political objective of the MU should instead be the promotion of human rights and liberty, and the solution of chronic problems such as the Israeli-Palestinian conflict.

The driving force essential to any progress is education and the ensuing scientific and technological development. General education will not only improve the well-being of society on all levels, but will also create an atmosphere of enlightenment that resists dogmatic and radical practices. Modern education will provide new skills and economies and the means for positive participation in the world market. And building a strong science base for research and development in the Mediterranean Basin, especially in the South, will have real impact with mutual benefits, allowing scientists and the rest of civil society to work together to alleviate many problems of significance to the region such as illegal immigration, illiteracy, food shortages, energy demands, water resources, climate change, infectious diseases, and the dearth of democratic governance. Moreover, creating such a base through sustainable cooperative programs with the North will limit brain drain and channel the energy of youth into a knowledge-based world economy.

The new MU initiative could turn into a historic milestone, building on the 1995 Euro-Mediterranean Partnership (the Barcelona Process), provided that there is a genuine desire for North-South support and partnership. Building education and the science base, bridging cultures through strong collaborative programs, and boosting economic and political benefits are the triad on which the MU should stand. These objectives will not see the light of day if the purpose of the MU is mainly political—rather, the focus should be "scientopolitical," a phrase coined here to emphasize the importance of education and science to the advancement of political and human affairs.

– Ahmed Zewail

10.1126/science.1164682

CHEMISTRY

Redshift Resolution

Dyes operate by absorbing specific wavelengths of light, thereby changing the overall color we perceive in looking at a dyed liquid or solid. In general, though, the color a dye confers depends not only on its molecular structure but on the medium in which it's been dissolved or suspended. Renger *et al.* derive a remarkably simple relation for predicting how the electronic absorption spectrum of an isolated nonpolar dye molecule will shift upon dissolution in a nonpolar solvent. In contrast to polar media, the influence of the weaker and more rapidly varying charge distributions in nonpolar media has remained puzzling. A prevailing model derived more than 50 years ago suggested that the absorption spectrum should shift to longer (redder) wavelengths as the strength of the absorption increased. However, the authors show that instead, the shift scales with the energy of the state accessed by the light absorption. They support their relation by accurately predicting spectral shifts for the bacteriochlorophyll a and bacteriopheophytin a chromophores. — JSY

Proc. Natl. Acad. Sci. U.S.A. **105**, 13235 (2008).

BIOMEDICINE

Folding Proteins

In many patients who suffer from loss-of-function genetic diseases, the missing protein is translated, but a point mutation causes misfolding and subsequent degradation of the protein. Pharmacologic chaperones, which help to restore function by binding to and stabilizing successfully folded proteins, have shown some therapeutic promise, but are inherently disease-specific. A more general strategy to encourage proper folding would be to enhance cellular protein homeostasis mechanisms, including the unfolded protein response (UPR) and the heat-shock response (HSR). Mu *et al.* have identified two small molecules, celastrol and the proteasome inhibitor MG-132, that each increase mutant protein folding and activity in patient-derived cell lines from two different lysosomal storage diseases, Gaucher and Tay-Sachs. These compounds up-regulate multiple UPR and HSR components, and two UPR proteins of the endoplasmic reticulum (IRE1 and PERK) are required to mediate the beneficial effects in both cell lines. Coapplication of these drugs with known pharmacological chaperones had a synergistic effect and increased the activity of the mutant proteins to at least 50% of wild-type activity. These results provide a proof-of-principle milestone in the therapeutic approach of developing protein homeostasis regulators to treat a range of loss-of-function diseases. — NM*

Cell **134**, 10.1016/j.cell.2008.06.037 (2008).

ECOLOGY

Have Your Plants and Eat Them, Too

An experimental study of Arctic vegetation shows that herbivores can exert a strong influence on the ecological outcomes of climate warming in plant communities. Numerous studies have shown that warming leads to changes in the biomass, structure, and composition of plant communities. In Arctic tundra in particular, warming leads to increases in aboveground plant biomass and of shrubby vegetation at the expense of grasses. Post and Pedersen conducted a 5-year experiment in which large vertebrate herbivores (musk ox and caribou) were either excluded from or allowed access to artificially warmed or ambient temperature exclosures. The warmed plots from which the herbivores had been excluded showed the expected transition to higher biomass and domination by woody plants (dwarf birch and willow). However, the plant communities on the grazed warmed plots were indistinguishable from those on the ungrazed ambient plots after 5 years. These results suggest that large herbivores might be useful in mitigating the effects of climate change in tundra—and perhaps in other rangeland habitats. — AMS

Proc. Natl. Acad. Sci. U.S.A. **105**, 12353 (2008).

APPLIED PHYSICS

An Umbrella or a Sieve?

Modern fabrication techniques can not only produce materials that show strong wetting or repulsion of water or organic liquids, but also offer dynamic control of the wetting behavior in certain cases. Lifton *et al.* fabricated a silicon membrane with honeycomb-shaped pores, overcoated with a nanonail architecture and an organic self-assembled monolayer or fluoropolymer. Under normal conditions, the membrane repels droplets of water or organic liquids, but after application of a voltage pulse, the droplets undergo an electrowetting transition and the

Continued on page 1421

*Nilah Monnier is a summer intern in *Science*'s editorial department.

CREDITS (TOP TO BOTTOM): © PURPLE MARBLES/ALAMY; ERIC POST

Continued from page 1419

fluid seeps into the pores. For a water droplet, to ensure that the fluid passed through the membrane, the authors placed a hydrophilic glass fiber filter underneath the membrane. They further fabricated a battery in which the tunable membrane keeps the liquid electrolyte and solid electrode separate until a voltage pulse is applied. Because there is no liquid penetration during storage, no electrochemical reactions occur, and so this sort of battery should have an extremely long shelf life. — MSL

Appl. Phys. Lett. **93**, 43112 (2008).

BIOCHEMISTRY

In Capsule Form

Intracellular membrane-bounded compartments—the mitochondrion, chloroplast, and nucleus—define the modern eukaryote. Bacteria make do without internal membranes, yet it is becoming evident that they do possess, nevertheless, intracellular nanosized environments. Sutter *et al.* describe the structure of the latest such oasis—an icosahedral shell 25 nm in diameter, formed by 60 monomers of the protein encapsulin. As one might intuit, encapsulin is structurally similar to viral capsid proteins, although any ancestral commonalities are no longer visible in their amino acid sequences.

Compartments can be useful for sequestering small molecules, either because they are valuable or because they might cause harm if allowed to diffuse. Encapsulin appears to offer both kinds of functions, because biochemical experiments identified docking sites for a peroxidase and for a ferritin-like protein, with the latter catalyzing the storage of iron as ferrihydrite and the former detoxifying potentially toxic oxygen species. — GJC

Nat. Struct. Mol. Biol. 10.1038/nsmb.1473 (2008).

SIGNAL TRANSDUCTION

Life Beyond Kinase

Phosphoinositide 3-kinase (PI3K) enzymes function by transducing the signals from receptor tyrosine kinases and heterotrimeric guanine nucleotide–binding protein (G protein)–coupled receptors. They catalyze the formation of lipid second messengers. Ciraolo *et al.* explored the biological role of p110β, one of three types of catalytic subunits that make up PI3Ks in mammalian cells. They created mice that express a mutant form of the enzyme that lacks kinase activity. Although complete loss of individual p110 subunits is deadly, the animals expressing the mutant p110β survived, indicating that p110β has functions that are independent of its kinase activity. The mutant animals grew more slowly than normal animals and, as they got older, showed signs of insulin resistance. They also showed impaired signaling from G protein–coupled receptors. In a breast cancer model induced by the ERBB2 oncogene, which signals through PI3K, the mutant animals showed fewer tumors, which grew at a reduced rate. In independent work, Jia *et al.* explored the effects of tissue-specific depletion of p110β in the liver and provide evidence for kinase-independent roles of p110β and for impaired insulin action in animals lacking p110β. They found that loss of p110β also decreased tumorigenesis in a mouse cancer model caused by loss of the lipid phosphatase PTEN. These studies indicate that tissue-specific actions of p110β may make it a potentially effective target for therapeutic regulation. — LBR

Sci. Signal. **1**, ra3 (2008); *Nature* **454**, 776 (2008).

PSYCHOLOGY

Socialized Learning

The ventral and dorsal neural streams that mediate the visual processing of objects have been described as specializing in what and where, respectively; that is, information about object identity flows through a different channel than that about object location. A related distinction applies to object perception for the purpose of action (where my car is parked) as contrasted with the purpose of recognition (my car is green). Yoon *et al.* demonstrate that pre-verbal infants (9 months old) form object representations that exhibit a similar dissociation between location and identity—and that the mode of information transmittal dictates the channel of reception. Introducing an explicitly social teaching context by having an actor point to an object while speaking to the infant biased the infant to remember the object's features such that a subsequent presentation of the same object at a second location did not evoke surprise (as assessed by looking time), whereas displaying a new object at the original location did. Conversely, reaching toward an object while verbalizing in an impersonal fashion primed the retention of where information rather than what. Adults, of course, have no difficulty in retaining and retrieving both types of representations, but the nascent neural processing capacities of infants appear to be influenced by social context as well as visual fundamentals. — GJC

Proc. Natl. Acad. Sci. U.S.A. **105**, 13690 (2008).

CREDIT: MARKUS SUTTER

www.sciencemag.org

Science

1200 New York Avenue, NW
Washington, DC 20005
Editorial: 202-326-6550, FAX 202-289-7562
News: 202-326-6581, FAX 202-371-9227
Bateman House, 82-88 Hills Road
Cambridge, UK CB2 1LQ
+44 (0) 1223 326500, FAX +44 (0) 1223 326501

SUBSCRIPTION SERVICES For change of address, missing issues, new orders and renewals, and payment questions: 866-434-AAAS (2227) or 202-326-6417, FAX 202-842-1065. Mailing addresses: AAAS, P.O. Box 96178, Washington, DC 20090-6178 or AAAS Member Services, 1200 New York Avenue, NW, Washington, DC 20005

INSTITUTIONAL SITE LICENSES please call 202-326-6755 for any questions or information

REPRINTS: Author Inquiries 800-635-7181
Commercial Inquiries 803-359-4578

PERMISSIONS 202-326-7074, FAX 202-682-0816

MEMBER BENEFITS AAAS/Barnes&Noble.com bookstore www.aaas.org/bn; AAAS Online Store http://www.aplsource.com/aaas/ code MKB6; AAAS Travels: Betchart Expeditions 800-252-4910; Apple Store www.apple/epgstore/aaas; Bank of America MasterCard 1-800-833-6262 priority code FAA3YU; Cold Spring Harbor Laboratory Press Publications www.cshlpress.com/affiliates/aaas.htm; GEICO Auto Insurance www.geico.com/landingpage/go51.htm?logo=17624; Hertz 800-654-2200 CDP#343457; Office Depot https://bsd.officedepot.com/portalLogin.do; Seabury & Smith Life Insurance 800-424-9883; Subaru VIP Program 202-326-6417; VIP Moving Services http://www.vipmayflower.com/domestic/index.html; Other Benefits: AAAS Member Services 202-326-6417 or www.aaasmember.org.

science_editors@aaas.org	(for general editorial queries)
science_letters@aas.org	(for queries about letters)
science_reviews@aaas.org	(for returning manuscript reviews)
science_bookrevs@aaas.org	(for book review queries)

Published by the American Association for the Advancement of Science (AAAS), *Science* serves its readers as a forum for the presentation and discussion of important issues related to the advancement of science, including the presentation of minority or conflicting points of view, rather than by publishing only material on which a consensus has been reached. Accordingly, all articles published in *Science*—including editorials, news and comment, and book reviews—are signed and reflect the individual views of the authors and not official points of view adopted by AAAS or the institutions with which the authors are affiliated.

AAAS was founded in 1848 and incorporated in 1874. Its mission is to advance science, engineering, and innovation throughout the world for the benefit of all people. The goals of the association are to: enhance communication among scientists, engineers, and the public; promote and defend the integrity of science and its use; strengthen support for the science and technology enterprise; provide a voice for science on societal issues; promote the responsible use of science in public policy; strengthen and diversify the science and technology workforce; foster education in science and technology for everyone; increase public engagement with science and technology; and advance international cooperation in science.

INFORMATION FOR AUTHORS
See pages 634 and 635 of the 1 February 2008 issue or access www.sciencemag.org/about/authors

More "Ötzis" in the Alps?

Glacier-free Schnidejoch and (*inset*) 5000-year-old piece of a leather shoe.

Receding Alpine glaciers have uncovered a trove of ancient artifacts in recent years. Last month, Swiss archaeologists announced that they had dated some of the items to as far back as 4500 B.C.E.—1000 years before the famous Iceman.

The owner of the items—a piece of wooden bowl and leather from a shoe—remains missing. But he has been named "Schnidi" after the Schnidejoch pass, where the items were found. "We now know that the findings at Schnidejoch are the oldest [yet discovered] in the Alps," said Albert Hafner, chief scientist at the Archaeological Survey of the Canton of Bern, at a news conference.

Since 2003, when record-high summer temperatures caused extensive melting of the ice at the 2756-meter-high pass, archaeologists have retrieved 300 items of hunting gear, fur, leather and woolen clothing, and tools belonging to early travelers or hunters moving between the Rhône Valley and parts north. Radiometric dating at the Swiss Federal Institute of Technology indicates that a bow, a birch-bark quiver, and arrows were dropped in the pass in the early Bronze Age, about 4000 years ago. Other finds include Roman coins and needles dating to about 200 C.E. and fragments of early and late Medieval apparel.

There is a wealth of data for climatologists as well, says University of Bern climatologist Martin Grosjean. "The findings allow us to accurately reconstruct glacier fluctuations in the Alpine area in prehistoric times," says Grosjean, who notes that periods of human passage at Schnidejoch nicely fit with periods when glaciers were in retreat and would have allowed travel.

CREDITS: (TOP AND INSET) ARCHÄOLOGISCHER DIENST DES KANTONS BERN; (RIGHT) JO YONG-MAK/REUTERS/LANDOV; (BOTTOM) © ARCANE-CENEG

Play Me a Molecule

Need to pry your kids off the games console? Tell them it has better things to do: molecular simulations.

Researchers in Spain have taken distributed computing to a new level. Instead of harnessing idle PCs for routine chores such as looking for signals from extraterrestrials, the new PS3Grid.net is exploiting the much greater capacities of idle Sony PlayStation3 (PS3) consoles. It's effectively "a new class of supercomputer," says project coordinator Gianni De Fabritiis of the University Pompeu Fabra in Barcelona.

Last year, De Fabritiis and colleagues developed software to exploit a new graphics chip that can process tens or hundreds of data streams in parallel, unlike the single fast stream in a normal PC. Not many such chips were available—until the launch of the PS3 in late 2006. The Barcelona team asked PS3 owners to donate their downtime. In the past year, about 400 machines have been signed up, which the team is using to perform complex biomolecular simulations. Plans are to let other research groups use it just as they would a conventional supercomputer.

Last month, the researchers launched a companion network for graphics processing units in normal PCs to add even more computing power. Once they reach their target of a 1000-machine network, they'll be able to do "accurate virtual screening of hundreds of molecules," De Fabritiis says.

Snuppy Dynasty Founded

Scientists reported this month that Snuppy, the world's first dog clone born in 2005, is now a daddy. His sperm was used to inseminate two cloned female fellow Afghan hounds. Of 10 puppies born in May, nine are still alive according to the dog-cloning team headed by Lee Byeong-chun at Seoul National University (SNU). Lee, originally on the dog-cloning team headed by disgraced researcher Hwang Woo-suk, was suspended by the university for 2 months in 2006 following the Hwang scandal.

Hwang and Lee are now on opposite sides of a patent fight over whether SNU or a U.S. company, BioArts International, has exclusive right to clone dogs for commercial purposes. Last week RNL Bio, the company Lee works with, announced that it had filed a patent infringement suit in Seoul against Suam Biotech, where Hwang works.

In Vitro Veritas

Enologists are going high-tech in their efforts to spot forged vintages. Besides using carbon-14 dating (see p. 1437), wine merchants may soon turn to particle accelerators to date old, valuable bottles of wine. Physicist Hervé Guégan, at the Centre d'Etudes Nucléaires de Bordeaux Gradignan in France, developed the technique of training low-energy protons on the bottle, causing its atoms to emit x-rays. The x-ray spectrum provides a kind of fingerprint that can be compared with a database to determine the age and provenance of the bottle. For example, French bottles made before 1957 typically contain traces of magnesium, so a bottle with a chromium signature dated before 1957 would be fake. The technique wouldn't be able to identify the exact year a wine was bottled, Guégan says, "because the chateau could have bought the bottle a year or two earlier."

Stephen Williams, CEO of the Antique Wine Co. in London, says his company is partnering with Guégan to develop the technology for commercial use. Every day it sells bottles valued between $1000 and $20,000. So "being sure about their age is … quite important to us."

Campaigns

DRIVEN. As a Swiss teenager, Louis Palmer dreamed of driving around the world in a solar-powered car. On 3 July 2007, the Lucerne native—a teacher by trade—set off in one built for $10,000 with help from businesses and universities across Switzerland. Last week, he cruised through Washington, D.C., and New York City having completed 80% of his 53,000-km journey.

The 3.5-m car, which can go 88 km/h, pulls a 5-m-long trailer covered in solar cells, and its passengers have included Rajendra Pachauri, chair of the Intergovernmental Panel on Climate Change, and New York City Mayor Michael Bloomberg. "I want to raise awareness that we can stop global warming and be independent from fossil fuels with today's technology," says Palmer, 36, whose next goal is to drive around the world in 80 days with a solar-powered car.

TWO CULTURES

SMASHING HIT. Not so sure what the Large Hadron Collider (LHC) will do? Then go to YouTube to view the Large Hadron Rap, a 4:49 music video in which science writer and jam master Katherine McAlpine explains it all as dancers gyrate gawkily in the tunnel that houses the CERN accelerator. McAlpine, 23, was a press contact for the U.S. contingent to the LHC. "I thought maybe we'll get a couple of thousand views," she says about the video, which has scored more than a million hits since it was posted 28 July.

Physicists are thrilled about the publicity. "[T]he text is way more accurate and to the point than most … articles about the machine and the experiments," writes blogger Roberto Corsini, an accelerator physicist at CERN (gravitasfreezone.wordpress.com). Watch out for more from McAlpine; she says the press office at the National Superconducting Cyclotron Laboratory at Michigan State University in East Lansing has asked her to rap about the lab's proposed accelerator.

CREDITS: (TOP TO BOTTOM): SOLARTAXI; JAY MAYFIELD/UNIVERSITY OF TENNESSEE; SOURCE: KATHERINE MCALPINE; EDDIE STANLEY

ON CAMPUS

MARRYING MINDS. Mathematical ecologist Louis Gross is already well-known as a scientist, teacher, and organizer. Now he will put those talents to use as director of a new institute at the University of Tennessee, Knoxville, that will tackle problems at the interface of mathematics and biology.

The National Institute for Mathematical and Biological Synthesis, launched last week, is modeled in part on a 13-year-old ecological synthesis center in Santa Barbara, California. Funded by a $16 million, 5-year grant that includes $11 million from the U.S. National Science Foundation and $5 million from the Department of Homeland Security, the institute will have a full-time staff of nine, about a dozen postdocs, and five core faculty members; it will also convene expert panels. Gross hopes to focus on animal diseases, such as the potential spread of pseudorabies from feral pigs in national parks to domestic swine, and explore other biological problems as diverse as wildfire control and cancer metastasis.

Gross, whose past experience includes developing a multi-scale model to help plan the restoration of the Florida Everglades, is "perfect" for the job, says animal-disease modeler Leslie Real of Emory University in Atlanta, Georgia. "He's very gifted at bringing biologists and mathematicians together."

Got a tip for this page? E-mail people@aaas.org

In the News >>

BIRD VERSUS MALL. Mark Anderson, 44, has studied the threats facing lesser flamingos for 18 years. But last month, he was suspended from his job with the South African provincial government apparently for trying to protect a new breeding site in a lake near Kimberley, some 475 km southwest of Johannesburg. Two of his scientific colleagues at the Department of Tourism, Environment and Conservation were also suspended.

The suspensions happened after a developer complained that Anderson and his colleagues—one of whom has also publicly supported a campaign to improve water quality in the lake—would not be neutral in reviewing the environmental impact assessment of a proposal to build shopping malls and more than 6400 houses on 382 hectares near the lake. Anderson, a staff ornithologist, says the 17 charges are trumped up. "It's remarkable that the department would go to this length," he says. Les Abrahams, the chief of staff for the department, which is also charged with economic development, won't comment on the charges but says it takes flamingo conservation seriously.

Anderson, who faces a disciplinary hearing starting on 15 September, already has a new job. Next month, he'll become executive director of the nonprofit advocacy group BirdLife South Africa.

Broad Institute's future secured 1432

Quantum flashlight 1433

Eyes down. All U.K. students take national exams at 16. Few study science subjects after that.

SCIENCE EDUCATION

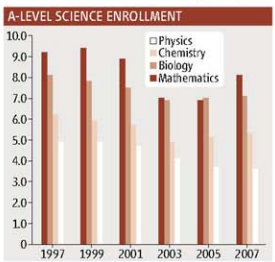
U.K. Education Reform: Too Much Of a Good Thing?

For more than a decade, the U.K. government has tweaked and revamped high school curricula and examination systems to stop a worrying slide in the number of children who study science and mathematics in their last 4 years at school. "The biggest problem is the high proportion of 15- and 16-year-olds who are dropping maths and science as soon as they possibly can," says Michael Reiss, director of education at the Royal Society.

Last week, the Royal Society issued a report that says the government implementation of science education reform is, well, unscientific. The changes have come so fast, one after another, that it's impossible to know whether anything has worked or just added to the problem, the report says. Moreover, new measures on the horizon, such as a high school science "diploma," are being rushed in without appropriate testing, the Royal Society warns. Curriculum reform, it concludes, should be managed by fully independent bodies, not politicians with short-term interests. "We strongly felt [reform] should be taken away from immediate political gain so as to get a more measured response," says polymer scientist Julia Higgins of Imperial College London, chair of the working group that produced the report.

The malaise in U.K. science education has been well-documented. With fewer pupils studying science at school, applications for some university science courses are going down. With fewer science graduates, the demand for them in industry is high and fewer go into teaching; schools then have trouble finding specialist science teachers and teaching standards drop; and, closing the circle, even fewer pupils study science. "The best learning experience comes from teachers who really know their subject," says Marianne Cutler, head of curriculum development with the Association for Science Education.

Just as in the United States, where each state manages its own education system, each of the four nations in the United Kingdom—England, Northern Ireland, Scotland, and Wales—has developed slightly different ways of teaching its children. However, certain features are common across the United Kingdom. At 14, pupils drop some subjects and typically continue with eight or nine through to age 16 when they take examinations in each subject, called Standard Grades in Scotland and GCSEs in the rest of the U.K. Then compulsory education ends but the majority continue their studies until 18, some taking vocational courses while those aiming at university take three or four courses leading to exams called A-levels ("highers" in Scotland).



Going down. The number of students starting A-level courses as a percentage of all 17-year-olds.

A major change in the teaching of the sciences in most of the United Kingdom came in the early 1990s, when combined science GCSE courses were offered in place of separate physics, chemistry, and biology courses to students under 16. Separate discipline courses were considered inflexible because students who didn't want to specialize in science post-16 might drop one or two of them, thereby getting a lopsided science education. Instead, all pupils broadly studied science but could choose between a more intensive course, the equivalent of two GCSEs, or a slimmed-down single science course.

The report says there is some evidence that this reform has reduced the number of children who pursue science after 16. In Scotland, where courses in the three basic sciences were retained, 12% of 17-year-olds studied for highers in physics in 2007, compared with just 3.6% studying for A-level physics in England. "The figures are strikingly awful if you look at them. It's a disaster," says Higgins.

In 2006, the English GCSE system was reorganized with the intensive double-science GCSE scrapped in favor of a range of single-award GCSEs, including separate physics, chemistry, and biology exams. The range of courses available is now "increasingly complex and targeted," the report says. Cutler believes the new GCSEs are going "in the right direction," but "they should have been piloted more widely and over a longer period."

The next step in the English school system, post-16 (the equivalent of junior and senior years at U.S. high schools), begins a makeover this year with the introduction of

CREDIT: R.J. VIEIRA/AAAS; SOURCE: DFES/DCSF/ONS/THE ROYAL SOCIETY

Bomb tests leave their mark on cells
1434

Space and the U.S. election
1438

diplomas. These alternatives to A-levels combine theoretical study with practical experience, including in the workplace. Diplomas launched this year include engineering, IT, and society, health, and development. The government plans to launch a science diploma in 2011. Meanwhile, Wales is testing a European-style baccalaureate, and Scotland is in the midst of a major review of all its curricula from ages 3 to 18.

The report does not say recent and upcoming curriculum reforms are misguided but rather that the U.K. nations have adopted a "nonscientific attitude to the introduction of change," Reiss says. He contrasts the education reforms to the care and meticulousness with which drugs are tested in clinical trials: "There's nothing remotely like that for the introduction of major change in education." He's particularly concerned about the upcoming science diploma in England. This "might fail precisely because it won't be piloted," he says.

The report concludes that the political pressure to deliver results before a government faces the next election is not compatible with methodical educational reform. That's why it called for an independent body to take charge of science education curricula. Reiss says they were open-minded about how such a body should be organized and constituted, but "we wanted it to be at arm's length from government."

The dwindling science pipeline feeding U.K. universities has had a noticeable impact: 22 physics departments have closed since 1997 (*Science*, 4 February 2005, p. 668), leaving fewer than half of U.K. universities now offering undergraduate physics degrees. Chemistry is in similar straits. Last December, the government commissioned a report on the state of physics in the U.K. from Bill Wakeham, vice chancellor of Southampton University. It is due to be completed this month.

–DANIEL CLERY

PALEOANTHROPOLOGY

Brainy Babies and Risky Births for Neandertals

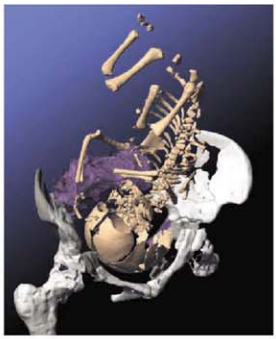
Any new mother can tell you that modern human infants have exceptionally large brains, which makes giving birth more difficult for us than for other primates. Now, a new study of a rare Neandertal newborn and two infants shows that our closest relatives were born with brains as large as ours and that those brains grew rapidly during the first few years of life.

This suggests that the uniquely human pattern of building big brains in utero and also expanding them quickly in infancy evolved long ago, before Neandertals and modern humans split from a common ancestor roughly half a million years ago, according to a report this week in the *Proceedings of the National Academy of Sciences*. Another part of the study, an analysis of a Neandertal woman's pelvis, shows that "Neandertal women had to face similar obstetric problems as modern human women," says co-author Christoph Zollikofer, a neurobiologist at the University of Zurich (UZ) in Switzerland.

As adults, the extinct Neandertals had brains and bodies larger than those of living people. But little has been known about their early brain development because few fossils have been found of Neandertal newborns or female pelves. A 1990 study of 10 Neandertal fossils between the ages of 2 and 10 found that their brain volumes were as large as those of modern humans. But the new study uses "amazing specimens" to provide the first data on infants, says Christopher Dean of University College London, who worked on the 1990 study.

Zollikofer and anthropologist Marcia Ponce de León of UZ took computed tomography scans of the most complete Neandertal neonate, which died 1 to 2 weeks after its birth about 63,000 to 73,000 years ago in Mezmaiskaya Cave in Crimea, Russia. They created a virtual reconstruction of the infant and calculated its brain size at birth to be 381 to 416 cubic centimeters (cm^3), within the range of modern humans. They also calculated brain size for two infants aged 19 months and 24 months from Dederiyeh Cave in Syria and found them to be at the upper end of the size range for modern human infants of those ages. That suggests that Neandertals' brains grew even faster during infancy than do those of modern humans. That conclusion "seems pretty robust," says paleoanthropologist Jay Kelley of the University of Illinois, Chicago, although he notes it is based on only one neonate.

The team also did a virtual reconstruction of the pelvis of a Neandertal female from Tabun Cave in Israel. Although the top of the pelvis was slightly wider than that of modern humans, they concluded that the anatomy of both Neandertals and modern humans limited fetal brain size to 400 cm^3. And in Neandertals, like modern humans, the large-brained fetuses must have rotated as they descended the birth canal, making birth difficult.



CREDIT: M. PONCE DE LEON AND C. ZOLLIKOFER/UNIVERSITY OF ZURICH

Bearing down. This computerized reconstruction of a Neandertal birth shows that Neandertal newborns had brains as large as those of modern humans.

Neandertal mothers would have spent a great deal of energy to fuel their offsprings' large and rapidly growing brains, note the authors. They propose that modern humans, whose brains and bodies have shrunk over the past 40,000 years, may thus have had an energetic advantage. Although provocative, that hypothesis is "the weakest part of the paper," says paleoanthropologist Jean-Jacques Hublin of the Max Planck Institute for Evolutionary Anthropology in Leipzig, Germany. He and Zollikofer agree that more data are needed to test whether it was smaller, more fuel-efficient brains that gave modern human mothers the evolutionary edge. –ANN GIBBONS

WORKFORCE

India Hopes New Fellowships Will Attract Expat Scientists

Synthetic chemist Gopalan Sampathkumar had always planned to return to his native India after finishing his postdoc at Johns Hopkins University. But a job offer last year from Arizona State University's Biodesign Institute seemed a lot more appealing than the prospect of earning $500 a month and battling a suffocating bureaucracy back home.

Last week, India's Department of Biotechnology (DBT) announced that it was teaming up with the Wellcome Trust, the U.K.-based biomedical charity, to level the playing field in a big way: a 5-year, $140 million program to support up to 375 scientists in all stages of their careers. "Our goal is to attract individuals who can go on to become leaders of India's growing biomedical research enterprise," says S. Natesh, a senior adviser to DBT.

The initiative has its roots in a DBT fellowship program for young scientists that helped persuade Sampathkumar to turn down the Arizona offer and join the National Institute of Immunology (NII) in New Delhi. That program, which offered twice the normal starting salary, received 72 applications for only 10 slots when it was launched in 2006 and convinced DBT officials that there was pent-up demand. "I still couldn't afford a mortgage on that salary. But at least it made family life in Delhi viable," says Sampathkumar, who returned last fall with his wife and two children.

The new program will provide annual salary packages ranging from $16,000 to $30,000 for 3 to 5 years, depending on experience. The awardees may work at Indian institutions of their choice. It does not include lab start-up packages, although recipients are free to negotiate one with their employer. They may also request funding for research materials in addition to the fellowship. Sampathkumar received a 2-year, $400,000 package from NII that has allowed him "to recreate the lab I worked at as a postdoc at Johns Hopkins University—we've bought exactly the same brand of equipment, from biosafety cabinets to fume hoods to incubators," he says. DBT also plans to establish several research facilities in the next 5 years, including one for stem cells and regenerative medicine and another for translational health research.

The new initiative builds on a smaller program run by the Wellcome Trust that currently provides fellowships to about 20 senior Indian researchers. "For a country of 1.2 billion, that's a grain of sand," says Natesh, who along with other Indian officials proposed that DBT team up with Wellcome. The charity's director, Mark Walport, says Wellcome sees the program as a way to help "establish a critical mass of Indian researchers who will jump-start the development of academic biomedicine" in the country.

CREDIT: AMINA KULSAR

Top of the line. Gopalan Sampathkumar says the equipment in his Indian lab is on a par with what he used in the United States.

The fellowships should attract talented scientists, Sampathkumar agrees. But that's not the only ingredient for a productive lab, he says. Unlike in the United States, "where I could pick up the phone and order supplies worth up to $2500," Sampathkumar says ordering anything at NII requires going through a purchase department that is often slow to respond. And a scientist might have to wait days for a trained technician to fix an important piece of equipment. "Every day that we lose because of purchasing or equipment hassles, we fall behind our competitors," he cautions.

Natesh says the DBT-Wellcome Alliance, an independent public trust that will administer the fellowships, will work with institutions to ensure that awardees get appropriate support and mentoring. The alliance plans to solicit applications later this year and hopes to announce the first round of awards—40 early-career fellowships for new Ph.D.s and 20 intermediate fellowships for those with postdoctoral experience—by mid-2009. Awards for 15 senior scientists, an extension of the existing Wellcome program, will be announced next month.

–YUDHIJIT BHATTACHARJEE

Stem Cell Article Retracted

Last week, *The Lancet* retracted a stem cell therapy paper without the consent of lead author Hannes Strasser, a urologist at the Medical University of Innsbruck. The move comes after an Austrian investigation uncovered ethical concerns about the conduct of the study. The paper, which reported results from an experimental stem cell therapy for urinary incontinence, was published in June 2007 along with an outside commentary hailing the findings as ushering in "a new era in urogynecology." The results of the Austrian investigation are confidential, but according to a *Lancet* comment last week, the report found that the study "was conducted neither according to Austrian law nor according to [international] standards. … The report found that there were critical deficiencies in the way patients' consent was obtained." In addition, *The Lancet* said, "the inspectors raise doubts as to whether a trial [conducted specifically] as described in *The Lancet* ever existed."

Strasser didn't respond to messages from *Science* seeking comment. In a June letter to *Nature*, he asserted that Austria's Ministry of Health had approved the study prior to its start. He also says that the medical school has offered the treatment to consenting patients outside the study after initial, positive results. After concerns about the paper surfaced, co-author Georg Bartsch of the Innsbruck urology department asked *The Lancet* to remove his name.

–RACHEL ZELKOWITZ

Cuba Law Struck Down

A 2006 Florida law banning state universities from sponsoring travel to neighboring Cuba and other countries under a U.S. trade embargo has been declared unconstitutional by a federal district court. The 28 August ruling allows Florida scientists to resume research in embargoed countries as long as they don't use state funds. The Florida Travel Act (*Science*, 9 June 2006, p. 1450) restricted the use of any funds—even from federal or private sources—by state-funded institutions. The American Civil Liberties Union challenged the law on behalf of Florida International University (FIU), and U.S. District Judge Patricia Seitz ruled that the law's restrictions infringed upon federal authority. The verdict should "make collaboration with Cuban colleagues much easier," says FIU geographer Jennifer Gebelein, who studies the impact of land-cover changes in Cuba on surrounding coral reefs.

–YUDHIJIT BHATTACHARJEE

PHILANTHROPY

Broad Gives $400 Million More to Cambridge Institute

The billionaire founder of a 10-year experiment in team science begun at the Massachusetts Institute of Technology (MIT) and Harvard University has decided halfway through that it's working so well it should be made permanent. Last week, Los Angeles businessman Eli Broad announced a $400 million gift that will allow the Broad Institute, already a genomics research powerhouse, to become a self-sustaining entity. "I think we've all agreed it's been a resounding success," Broad told reporters.

Broad donated $100 million in 2003 to create the institute after visiting Eric Lander's huge lab at the MIT-affiliated Whitehead Institute for Biomedical Research as it was winding up its part in sequencing the human genome. The new institute's mission was to move genomics into the clinic (*Science*, 20 June 2003, p. 1856). It was set up administratively as part of MIT, with Lander and three other scientific stars from MIT and Harvard as its founding faculty. It has since attracted another $100 million from Broad and $100 million from the Stanley Medical Research Institute for research on the genetics of psychiatric diseases.

Broad Institute researchers have played prominent roles in projects such as the HapMap, which studied human genetic diversity; a consortium to develop RNAi research tools; and a search for mutations in human cancers. The $150-million-a-year institute now has about 1100 full- and part-time permanent staff and 118 affiliated faculty members. Lander says its two strengths are strong technology and a structure that allows it to "self-assemble" teams from MIT and Harvard, including its 17 affiliated hospitals. "It is a really good, innovative model," says Bruce Stillman, president of Cold Spring Harbor Laboratory in New York.

Proof of concept. Billionaire Eli Broad discusses his latest gift to the Broad Institute, directed by Eric Lander (*left*).

Broad's latest gift will allow the institute to have something it has never had—an endowment. "[It] will secure the permanency of the institute," says Broad, who hopes other gifts will raise the pot to $1 billion. Although the institute will still be governed by a board drawn from MIT and Harvard, its standalone status will give it greater flexibility in paying its scientists, staffers say.

–JOCELYN KAISER

U.S. ELECTION

McCain, Obama Present Their Wars on Cancer

With the U.S. presidential election less than 2 months away, both candidates explained last week how their Administrations would combat cancer. Appearing on a celebrity-studded television fundraiser, Stand Up 2 Cancer, that aired on 5 September, Republican John McCain and Democrat Barack Obama advocated somewhat different strategies but agreed on the need for better access to early detection technologies and more preventive care.

McCain's statement highlights legislation he supported in 2001 to improve access to clinical trials and, last year, to fund research on the environmental risk factors of breast cancer, a bill Obama endorsed as well. McCain also referred to his past support for doubling the National Institutes of Health (NIH) budget over 6 years, adding that "as President, [I] will make sure that our researchers have necessary funding to defeat cancer once and for all."

Obama offered a denser, arguably more detailed plan, which included doubling the budget for cancer research in 5 years, mainly through the National Cancer Institute, and boosting from about 4% to 10% the number of adults with cancer participating in clinical trials. He also said he would provide "additional funding for research on rare cancers and those without effective treatment options" and for the study of genetic factors driving cancer and outcomes.

"He's been hearing from scientists ... who have told him that we're stagnating" because of a flat NIH budget, says Neera Tanden, an Obama domestic policy adviser. Tanden adds that "there's no reason to assume" NIH, which enjoyed a rapid doubling of its budget in the late 1990s, would suffer a second crash landing of the kind it's experiencing today if its budget again rose dramatically. Tyler Jacks, director of the Koch Institute for Integrative Cancer Research at the Massachusetts Institute of Technology in Cambridge, says he applauds McCain's pledge to better coordinate public and privately funded research because foundations and individuals have greatly increased their spending on cancer. But he found it "a little odd" that McCain emphasized past legislation rather than looking ahead.

Richard Marchase, president of the Federation of American Societies for Experimental Biology in Bethesda, Maryland, welcomes the willingness of both candidates to consider boosting the science research budget. But he cautions that "we're going to have much better results if we have a broader base than just cancer funding." Focusing too heavily on one disease, he says, could blunt the impact of "serendipity" in the lab.

–JENNIFER COUZIN

CREDIT: JOSH REYNOLDS/AP

PHYSICS

Quantum Flashlight Pierces the Darkness With a Few Percent as Many Photons

Anyone who has stubbed a toe while creeping through a darkened bedroom probably has wished to see better in the dark. Now, a theoretical analysis on page 1463 shows how to do that using strangely interconnected particles of light, or photons. The approach greatly reduces the number of photons needed to detect an object. It could be used for imaging samples in the lab or for spotting satellites in the skies, says Seth Lloyd, the paper's author and a self-described quantum-mechanical engineer at the Massachusetts Institute of Technology in Cambridge.

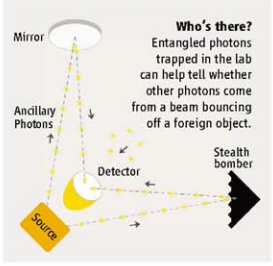
"This is a new way of exploiting quantum mechanics," says Giacomo Mauro D'Ariano, a theorist at the University of Pavia, Italy. Juan Perez Torres, an experimenter at the Institute of Photonic Sciences and the Polytechnic University of Catalonia in Barcelona, Spain, says putting the plan into practice should be "not impossible to do."

To tell if there's something out there, you could shine a search beam to see if it reflects back to you. However, an object may reflect only a tiny fraction of light in the direction from which it came, and the surrounding environment can generate photons itself. So you'll have to send out enough photons to determine if the number coming from a particular direction exceeds the number expected from the environment alone.

Quantum weirdness can whack down the number needed, Lloyd says. Suppose the photons in the beam are emitted one by one, and each can have any of 30 slightly different colors or frequencies. Then, bizarrely, quantum theory allows each photon to have all 30 different frequencies at once—although that delicate "superposition" collapses randomly to one frequency or another as soon as the photon is measured.

Moreover, two photons can be "entangled" so that each is in such an indefinite split-personality state, yet their frequencies are locked together. For example, the photons can be entangled so that when they are measured, the sum of their frequencies will equal twice the average of the 30 frequencies. Lloyd envisions sending one photon from each pair in the beam while keeping the other "ancilla" in a delay loop (see figure, above).

A photon arriving from the chosen direction is then compared with the ancilla. If the arriving photon started out in the beam, its frequency and the ancilla's will add to the correct sum. A photon arriving from the background has only a 1-in-30 chance of making a correct tally, so it is harder for background photons to masquerade as returnees from the beam. That difference makes it possible to detect an object with only 1/10 as many photons, Lloyd calculates. He says it should be possible to use 1 million frequencies in the future.



CREDIT: J. NEWFIELD/SCIENCE

Ironically, although the entanglement helps reveal the object through the background noise, the noise obliterates the entanglement. Without noise, the frequencies of every pair would add to the same sum. Mix in enough noise, and the sum varies from pair to pair essentially at random, signifying a complete lack of entanglement. Yet merely starting with entangled pairs is enough to improve the efficiency for detecting an object in the noisy case. "Any entanglement you put into the system is completely gone by the time you make the measurement," Lloyd says. "Nonetheless, it helps."

That's surprising, says experimentalist Torres—so much so that he wonders if entanglement is really necessary after all. "It might be that the entanglement is destroyed but that you still have certain [weaker] correlations that give you an advantage," he says. "But then it's not the entanglement that's producing the effect."

But the particular formula for the improvement suggests that entanglement, although destroyed, is key, says theorist D'Ariano. "There are many things to check, but this idea is very promising," he says. Lloyd says he hopes someone will do a proof-of-principle experiment within a year. **–ADRIAN CHO**

California to Overhaul Chemical Regulations

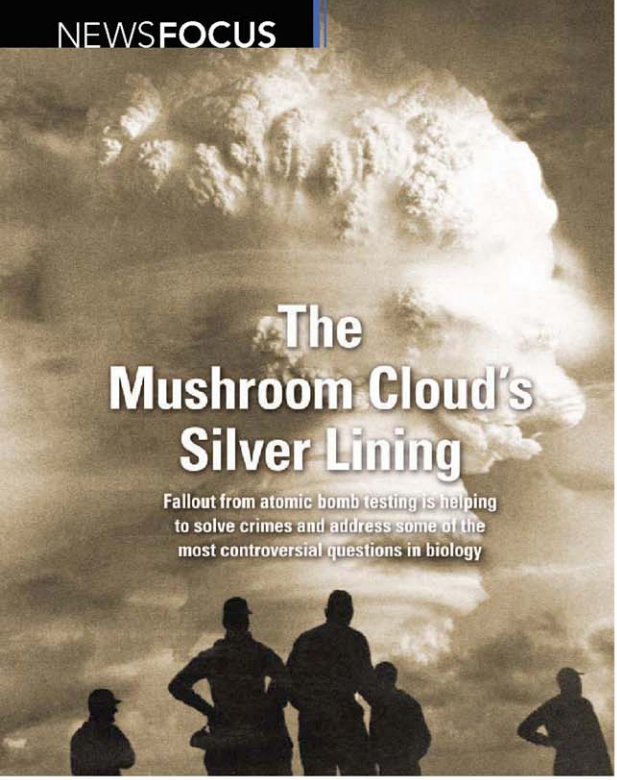
California Governor Arnold Schwarzenegger is expected to sign two bills passed last month that inject more science into the regulation of chemicals used in consumer products. But some environmentalists fear that the additional analysis might slow the new procedures.

California has traditionally addressed new threats on a chemical-by-chemical basis, most recently by banning phthalates in children's toys. In contrast, AB 1879 would give the Department of Toxic Substances Control (DTSC) authority to list and regulate chemicals of concern based on the extent of exposure and the risk posed to children and infants, with input from an advisory panel that includes scientists. Before it could regulate a chemical, however, DTSC would have to analyze the risk posed by the chemical through its life cycle from production to disposal, the safety of possible substitute chemicals, and the cost of implementing any rules. "I worry that this might be a bottleneck," says Richard Denison of the Environmental Defense Fund.

The second bill, SB 509, would create an online clearinghouse on chemical hazards. Both bills enjoy considerable support from industry, and the Pew Environment Group would like to see similar legislation adopted by the U.S. Congress. **–ERIK STOKSTAD**

A Matter of Degrees

Preliminary results from a 7-year project to improve Ph.D. completion rates at U.S. and Canadian universities indicate that whites, men, and international students are more likely to complete their degrees than women, other ethnic groups, and domestic students. That's what experts have long suspected. But there are also some surprising differences, according to a report this month from the Council of Graduate Schools. African-Americans have the greatest variance in completion rates by discipline, for example, although the numbers are too small to be statistically significant. Although 60% complete life sciences degrees in a 10-year period (the same as for whites), only 37% do so in math and the physical sciences. The project, funded by Pfizer Inc. and the Ford Foundation, supports additional data analysis as well as a range of interventions by 29 institutions—from additional mentoring to increased research opportunities—aimed at helping more students complete their degrees. **–JEFFREY MERVIS**



The Mushroom Cloud's Silver Lining

Fallout from atomic bomb testing is helping to solve crimes and address some of the most controversial questions in biology

Bomb boom. Hundreds of aboveground nuclear tests, like this one carried out in the Pacific in 1958, seeded the atmosphere with excess ^{14}C.

CREDIT: CORBIS

THE TWO MUMMIFIED BODIES IN THE Vienna apartment told a sad tale. The reclusive elderly sisters had clearly been dead for several years, but no one had noticed; neighbors in the upper-middle-class complex believed they had merely moved away. Stale bank accounts finally tipped off the police, who discovered the remains in December 1992.

Investigators found no evidence of foul play, so they focused on the question of who died first. Both sisters had large pensions and separate life insurance policies, and the insurance company of the woman who died last would collect the bulk of the funds. "There was a lot of money at stake," says Walter Kutschera, a physicist at the Vienna Environmental Research Accelerator at the University of Vienna in Austria. Not long after the bodies were found, a scientist from the university's forensics department approached Kutschera and his colleague, Eva Maria Wild, to ask if they could help crack the case. The forensics expert knew the pair had been using radiocarbon dating to determine the age of archaeological samples, and he wondered if the same technique could shed light on the year each sister had died.

It couldn't. Radiocarbon dating is a blunt instrument that relies on the slow decay of a form of carbon known as carbon-14 (^{14}C), which is incorporated into animals during their lifetime. The method works well for samples that are tens of thousands of years old, but it's only accurate to within a few hundred years.

Wild and Kutschera had another idea. Aboveground testing of nuclear weapons after World War II had injected ^{14}C into Earth's atmosphere, creating an abnormally high level of the isotope that has been tapering off since then. If the researchers could measure the amount of ^{14}C in something carbon-based that the sisters had generated just before death—fats in the bone, for example—and compare it with historic levels of ^{14}C in the atmosphere, they should be able to tell which year each sister expired.

It worked. Wild and Kutschera found that one sister had died in 1988 and the other in 1989. "One sister lived for some time next to the dead one," says Wild. Investigators closed the case, and Wild and Kutschera returned to dating ancient bones and seeds. But it would soon become clear that the "bomb pulse" technique had much more to offer. In the past decade, thanks largely to the pioneering work of an Australian postdoc with a taste for trying new things, groups have begun using the strategy for diverse causes such as identifying disaster victims, authenticating wine vintages, and tackling some of the most controversial questions in biology, including whether the human brain generates neurons throughout life.

Online
sciencemag.org
Podcast interview with the author of this article.

From pet shop to slaughterhouse

2001 started well for Kirsty Spalding, but by the end of the year she would be knee-deep in a failing project. The 29-year-old had just finished her graduate work in neuroscience at the University of Western Australia in Perth, and she was planning on spending a year in Europe as a postdoc before moving to the United States. On her way to interview at a couple of prospective labs at the Karolinska Institute in Stockholm, Sweden, Spalding caught a talk by Jonas Frisén, a prominent stem cell researcher there. "It wasn't what I had planned on doing," says Spalding, referring to Frisén's work on the formation of new neurons in the brain—a process called neurogenesis. "But I found him very personable and the work very interesting."

A few months later, Spalding was in Frisén's lab, trying to map neurogenesis in the zebrafish brain. But neither she nor her labmates had worked with the animal before, and they weren't aware that technical suppliers provided fish specially bred for laboratory study. Instead, Spalding biked over to a local pet shop and brought a few zebrafish back to the lab. Needless to say, the experiments didn't work.

Her mentor didn't lose faith, however. "I could tell that Kirsty liked challenges and that she was extremely entrepreneurial," says

Frisén. That made her perfect for a new project he had in mind. Familiar with the bomb-pulse work done by Wild and Kutschera, Frisén wondered if it could be applied to DNA. When a cell divides, ^{14}C in the environment is incorporated in new chromosomes, and thus the DNA effectively takes a snapshot of the amount of atmospheric ^{14}C—and hence the birth date—of the cell. If Frisén could exploit this, he might be able to show whether humans generate new brain cells throughout life—a central question in neuroscience. But no one would take on the project. Postdoc after postdoc turned him down, calling the work too risky and too difficult. When Frisén saw Spalding with the zebrafish, he knew he had found someone who wouldn't be daunted.

Spalding agreed. "I liked the problem-solving aspect of it, and I didn't have the burden of knowledge to know how difficult it would be," she laughs. Spalding's planned 1-year sojourn in Europe suddenly became an indefinite commitment.

To address neurogenesis in humans, Spalding needed brains from an animal with a similar life span, so she turned to horses, which can live more than 25 years. That meant trips to the local slaughterhouse. "I would watch them walk the horse in, … and then they would chop off its head and hand it to me," recalls Spalding, who had to excavate the skulls herself. "It's not so easy to hack your way into a horse's head. … It was not pretty."

Brains in hand, Spalding still had challenges to overcome, such as measuring a scarce isotope. ^{14}C makes up only one part per trillion of all of the carbon in the atmosphere. Most comes from cosmic ray collisions with nitrogen, but when the United States, the former Soviet Union, and other nations detonated more than 500 nuclear warheads aboveground in the 1950s and '60s, the atmospheric ^{14}C level doubled. It only began to dissipate when the Limited Test Ban Treaty of 1963 moved atomic tests underground (see illustration).

Despite these elevated atmospheric concentrations, only about one atom of ^{14}C incorporates into every 15 cells. So relatively huge amounts of tissue—up to 5 grams, depending on the part of the body it comes from—are needed for even the world's most powerful isotope detectors to spot it. Horse brains were big enough to provide that amount, but Spalding also had to find a way to sift through a custard of fat, glia, and fibroblasts for the neurons she needed. After taking nearly a year to develop a technique, she was ready to pin ages on neurons and enter the ongoing fray over neurogenesis.

> *"The clinical implications are huge. … There are hundreds of great biological questions that can be answered [with this technique]."*
>
> —YUVAL DOR, THE HEBREW UNIVERSITY

The brain war

Pasko Rakic is a five-star general in a conflict that's been raging for more than a decade in the neuroscience field. The Yale University neuroscientist, who did pioneering work in how the primate brain forms, has famously established the beachhead position that the human cerebral cortex—a region key for memory, language, and consciousness—does not make new neurons after development. He's often made the point that such adult neurogenesis would be counterproductive, disrupting already formed memories, for example.

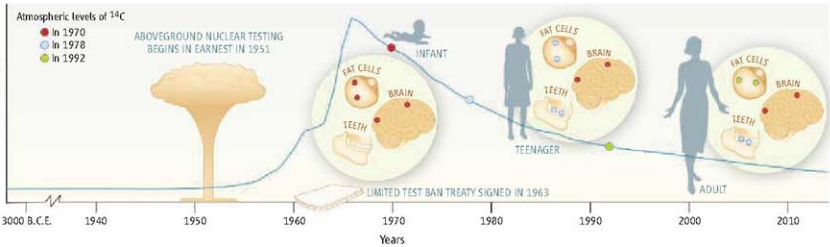
But in 1998, a research team found evidence to the contrary. It gave people with terminal cancer a synthetic compound called bromodeoxyuridine (BrdU), which inserts into newly synthesized DNA and thus serves as a marker for new cells. The compound was supposed to gauge tumor growth, but it also showed up in the hippocampus, the brain's learning and memory center (*Science*, 6 November 1998, p. 1018). A year later, Princeton University neuroscientist Elizabeth Gould bolstered the case for ongoing neurogenesis in the brain by giving adult macaques BrdU and finding it in the neocortex, a region responsible for language and consciousness in humans. But 2 years after that, Rakic injected a different DNA marker into monkeys and saw no new neurons in the adult brain. The field has been divided ever since.

"It's been extremely difficult to get any information in humans," says Gerd Kempermann, a neurogenesis expert at the Center for Regenerative Therapies Dresden in Germany. BrdU is toxic, so it can't be given to healthy people, and Rakic has expressed concern that the compound confuses cells into dividing, leading to false positives.

^{14}C doesn't have that problem. It's not toxic, and like it or not, we've all absorbed it. "All of humanity is labeled," as Kutschera puts it.

As the salvos continued in the neurogenesis debate, Spalding had proved that she could use the bomb-pulse technique to date brain cells in horses. She shipped her first human samples—from the brain's visual center, the occipital cortex—to Bruce Buchholz, who runs an isotope detector the size of a basketball court at Lawrence Livermore National Laboratory in California. Although nonhuman studies had suggested that the occipital cortex was a hotbed of neurogenesis, the ^{14}C data collected by Buchholz indicated that human neurons from this region had the same birth date as the people they came from. That meant no new visual neurons for adults. A year later, Spalding and colleagues found similar results in the human neocortex.

"It's really extraordinary work, and it's extremely clever," says Kempermann. "I think many people will take it as the final



CREDIT: P. HUEY/SCIENCE

Atomic child. ^{14}C levels in DNA from visual and memory neurons stay the same throughout life, indicating no neurogenesis in these brain regions. Levels in fat cells change, suggesting constant fat cell turnover. ^{14}C levels in tooth enamel remain constant and can be used to calculate a person's birth date.

word in the debate." Still, Gould notes that other regions of the human brain—such as the hippocampus—have yet to be tested with the technique. And she says that because the bomb-pulse method doesn't target individual cells, it may not be sensitive enough to pick up a small population of neurons that does divide and could contribute to repair and learning. Spalding was in the midst of addressing those questions when disaster struck a continent away.

CSI: Sweden

"Total chaos." That's how Stockholm's former chief medical examiner, Henrik Druid, describes the scene as bodies piled up at the Karolinska Institute morgue in the wake of the 2004 Indian Ocean tsunami that killed more than 200,000 people, including more than 500 Swedish tourists. "The bodies were so badly decomposed, you couldn't tell the teenagers from the old people," he says.

Hoping to help, Spalding approached Druid with some intriguing findings from her days at the slaughterhouse. In addition to analyzing horses' brains, she had looked at their teeth, showing that because enamel is permanent and forms early, its ^{14}C levels give an accurate estimate of the animal's age. Spalding asked Druid if the technique might be useful to him.

"At first I was skeptical," he recalls. But Druid didn't have many options. In the confusion surrounding the disaster, identifying materials such as x-rays and DNA samples from relatives had not been shipped with the bodies. "If you have no clue to the identity of a person, age and sex are the most important way to limit the search," he says. Anthropologists are only accurate to within about 10 years when trying to determine age from a skeleton. So, aided by Spalding, Druid applied the bomb-pulse technique to the teeth of six tsunami victims. After adding the time it takes for human enamel to form (about 12 years for wisdom teeth, for example), they were able to predict the ages of every victim to within 1.6 years, as borne out by the identifying materials that eventually arrived at Karolinska.

With further refinement, Druid has shaved the accuracy down to 1 year, and he's now using the approach to help Swedish investigators crack two unsolved homicides. "This is going to be very, very valuable for criminal investigation," says Druid. "In a year or two, you're going to begin seeing cases in the newspaper that were solved with this method." Spalding too has begun working with Swedish police—as well as with investigators in Canada—and she eventually hopes to set up a company to perform the tooth analysis. In preparation, she has taken business classes at night, all while forging ahead with her brain work—and a new project that would send her spinning in an entirely different direction.

> *"In a year or two, you're going to begin seeing [criminal] cases in the newspaper that were solved with this method."*
>
> —HENRIK DRUID, KAROLINSKA INSTITUTE

The fat offensive

In 2005, Spalding was presenting her brain findings at Karolinska when a member of the audience approached her. "A Ph.D. student came up to me and said he thought the ^{14}C work was something his dad would be interested in," she says. The father—a prominent researcher at Karolinska named Peter Arner—was grappling with a debate not unlike the one faced by the neurogenesis community.

This time the issue was fat. "If you go to any textbook, it will tell you that once a fat cell is born, you've got it forever," explains John Prins, an expert on fat-cell turnover at the University of Queensland in Brisbane, Australia. But there were some who believed that the blubber on our bellies and hips is constantly dying and being replenished. It's not just an academic debate: If you can make the body destroy more fat than it creates, you've got a ticket to weight loss.

But no one could conclusively address the question. "The techniques we have for measuring fat turnover are insufficiently sensitive and fairly inaccurate," says Prins. The best researchers could do was have volunteers drink heavy water, which contains elevated levels of an isotope of hydrogen known as deuterium, and look for that isotope in fat cells. "Not too many people want to drink heavy water," Prins says.

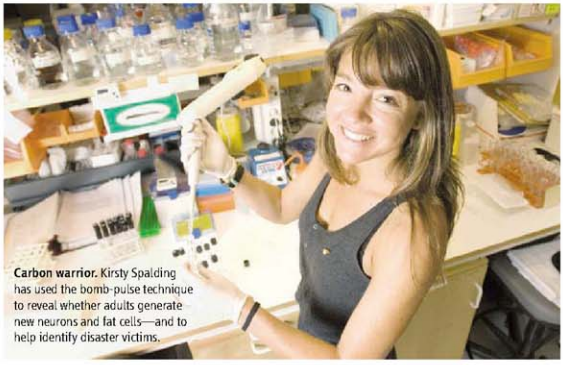
Spalding began working with Arner, and by 2006 she had developed a regimen for isolating fat cells from the vast array of other cells found in human flab. Analyzing fat biopsies and liposuction leftovers from people of various ages, Spalding showed that people born a few years before atomic bomb testing began had fat cells with high levels of ^{14}C, which only made sense if these cells were generated after the fallout had spiked the isotope's levels. When Spalding looked at people born after the bomb tests, she saw fat cells with different amounts of ^{14}C, levels corresponding to various dates on the bomb-pulse curve. In all, the data indicate that people replace half of their fat cells about every 8 years, Spalding reported this summer in *Nature*.

"It's a landmark paper … and a phenomenal advance on a number of fronts," says Prins. "You've got this technique out of *Star Trek*, and now everybody thinks that fat is a dynamic organ." No drug company would have looked into fat turnover before, he says, "but now people will start to consider therapeutic perspectives."

Loving the bomb

As the years go by, the ^{14}C level in the atmosphere is slowly returning to its prewar levels. Rising carbon dioxide emissions, chock-full of ^{12}C, have only hastened the isotope's demise. And yet the bomb-pulse technique is just taking off.

Both Spalding, who left Frisén's lab in



Carbon warrior. Kirsty Spalding has used the bomb-pulse technique to reveal whether adults generate new neurons and fat cells—and to help identify disaster victims.

CREDIT: STEFAN ZINVERMAN

FORGERS FACE THE NUCLEAR OPTION

Graham Jones knows a good wine when he tastes one. For nearly 2 decades, the enologist has been teaching students and winemakers at the University of Adelaide in Australia how to become better connoisseurs of the beverage—and how to spot a fake. So when the Australian wine industry became concerned about its vintage wines being disputed in the wake of surging exports to Europe in the late 1990s, it turned to Jones.

"We wanted to develop a technique that, if our wines were challenged, we had the ability to authenticate them ourselves," says Jones. A colleague suggested he take a look at the bomb-pulse technique (see main text). Applied to wines, the method should allow researchers to verify the year a wine was made. That's because, when grapes grow, vine leaves take up ^{14}C-containing carbon dioxide from the atmosphere and convert it to sugar, which eventually becomes the alcohol in wine. Jones's lab developed a procedure to separate the alcohol from other components of wine, but he needed a way to measure its ^{14}C content. So he turned to scientists at Australia's only nuclear research accelerator, based at the Australian Nuclear Science and Technology Organisation (ANSTO) in New South Wales.

A good year. Graham Jones has used the bomb-pulse technique to authenticate wine vintages.

The team was able to accurately calculate the vintage, within 1 year, of a variety of South Australian Cabernet Sauvignons bottled between 1958 and 1997. Although the technique is too expensive to be used regularly, Jones thinks the study warded off European regulators. "The fact that they haven't challenged any of our wines yet is a definite plus for the work," he says.

The ANSTO team has since moved from wine to illicit drugs. When sold illegally, narcotics like morphine tend to be produced and shipped quickly, says Jones, whereas legal morphine can sit around for a while after it's made. ANSTO researchers have shown that it's possible to date these drugs—via the ^{14}C content of the poppy plants they come from—as a way of gauging their legality.

Scientists elsewhere have targeted another type of illicit activity: poaching. Bruce Buchholz, who runs an isotope detector at Lawrence Livermore National Laboratory in California, is collaborating with researchers to date ivory tusks and lion teeth. Because tusks grow throughout an elephant's life, scientists can determine if one has the ^{14}C signature of a time after an ivory ban went into effect. Similarly, ^{14}C in teeth could ostensibly reveal whether hunters are killing off too many young male lions. Buchholz has also heard about groups using the technique to gauge whether a painting supposedly made before the 1940s is a recent forgery, based on the ^{14}C content of the canvas. "If it's supposed to be old and it has bomb carbon in it," says Buchholz, "you know something's wrong." **–D.G.**

2006 to become an assistant professor at Karolinska, and Frisén are expanding its applications. Entering debates similar to the ones about neurogenesis and fat turnover, they're looking at whether heart cells and insulin-producing beta cells in the pancreas renew throughout life or whether we're stuck with the ones we're born with. In tissues in which stem cells have been identified, they plan to examine how often these cells divide and how they are made.

"The clinical implications are huge," says Yuval Dor, a cell biologist at The Hebrew University–Hadassah Medical School in Jerusalem, Israel, and an observer of the bomb-pulse technique. "There are hundreds of great biological questions that can be answered. … We're all very much looking forward to how this will turn out."

The weight isn't all on Spalding and Frisén's shoulders. Other groups have begun to experiment with the technique as well. Like Frisén, diabetologists David Harlan and Shira Perl of the U.S. National Institute of Diabetes and Digestive and Kidney Diseases in Bethesda, Maryland, are using ^{14}C to measure turnover in beta cells. And Lawrence Livermore's Buchholz says he's been approached by a number of labs interested in everything from climate modeling (changing weather patterns are reflected in ^{14}C levels in coral) to dating confiscated ivory tusks and authenticating wine vintages (see sidebar, above).

Still, it's not a technique that most labs have the resources to adopt. "There are no kits you can buy to do this," says Buchholz. And most labs don't have access to the powerful isotope detectors needed to perform the ^{14}C analysis.

Critics also point out that the bomb-pulse technique has limitations. Although Spalding's work supported Rakic's stance on neurogenesis, Rakic notes that when damaged cells repair DNA, that DNA could incorporate new ^{14}C, suggesting new cell formation when there is none. Conversely, fat-turnover expert Prins says that new cells sometimes recycle DNA from dead cells, giving the impression—under ^{14}C analysis—that no new cells have been made.

And Spalding admits that the forensics applications have a shelf life: As ^{14}C levels recede to background in the atmosphere—Buchholz estimates a return to prebomb conditions by 2020—it will become harder and harder to tell a corpse's year of death. But she's optimistic that as isotope detectors become more sensitive—she's working with Wild and Kutschera to help make this happen—police will be solving cases with the technique for years to come. Brain, fat, and other clinical research won't be affected by the dissipation, as scientists can turn to tissue samples banked over the decades after the bomb tests.

Back at Karolinska, Spalding, Frisén, and a few other collaborators have just formed a Center of Excellence to map the regenerative potential of the entire human body. Over the next 10 years, they'll try to gauge the turnover of every cell type they can. "I love this technique," says Frisén. "We're having a lot of fun with it."

Next year, for a sabbatical, Spalding will head off to California, where she will look for new challenges while continuing her brain and fat research. Stay tuned for an upcoming paper on neurogenesis in the hippocampus—and some more surprises with fat turnover.

Meanwhile, at the birthplace of the atomic bomb in New Mexico, retired Los Alamos National Laboratory scientist Donald Barr reflects on what Spalding and the other bomb-pulse researchers are doing. He's been at the lab for more than 50 years, keeping tabs on nuclear fallout in the atmosphere, and he still comes in a couple of days a week to chat isotopes with his former colleagues. The mushroom clouds from nuclear detonations do indeed have a silver lining, he says. "There are questions we can now answer because of that testing that scientists never thought about at the time." **–DAVID GRIMM**

CREDIT: STEPHEN CLARKE

ELECTION 2008

Obama and McCain Are Swept Up In a Surprising Space Race

Space policy may not be on the minds of most Americans, but it's become an important issue in the race for the White House. How did that happen, and what does it mean for President Bush's 2004 vision for exploration?

Science and the 2008 Campaign

Stem cells, climate change, energy research, the teaching of evolution—these are today's hot-button science and technology issues. But in the contentious U.S. presidential race, the human space exploration program stole the limelight last month.

A savvy group of business boosters in electoral-vote–rich Florida and a small band of determined space advocates have convinced the Republican contender, Arizona Senator John McCain, and his Democratic rival, Illinois Senator Barack Obama, that NASA's fortunes are intertwined with their quest for the Oval Office. Vying last month to prove their space-friendly credentials, the two men visited the area around NASA's Kennedy Space Center in Florida, issued dueling policy statements, and insisted that they were eager to boldly go where humans have not been since geologist Harrison Schmitt closed the hatch on the lunar module in 1972. Returning to the moon even made it into the Republican Party platform finalized last week in St. Paul, Minnesota.

The impetus for the debate is the job losses connected to a 2010 phaseout of the aging space shuttle. Both candidates say they will consider postponing that retirement date while pushing for a new launcher that could speed humans to the moon by 2020. Both also want to bolster scientific research aboard the international space station still under construction—and question the Bush Administration's decision to mothball it in 2016.

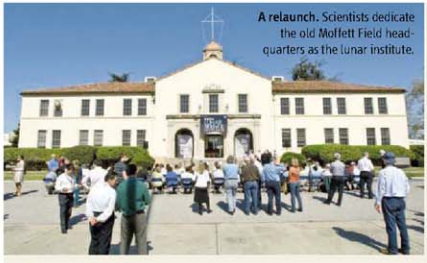
The unusual bout of political one-upmanship has broadened the debate over the agency's future beyond its traditional audience of university researchers and aerospace engineers who benefit from NASA's annual $17 billion largesse and their congressional supporters. It also promises to brighten the agency's current gloomy fiscal picture. "Raising the profile of space as a campaign issue in Florida is an excellent way to increase the budget of NASA," says Dale Ketcham, director of the Spaceport Research and Technology Institute, a consortium based at Kennedy that backs research and commercialization efforts. The debate provides a rare glimpse into how politics, economics, and science and technology interact to make a campaign issue. "It's lucky we're a swing area in a swing state," says Lynda Weatherman, president and chief executive officer of the Economic Development Commission of Florida's Space Coast, which played a starring role in placing space on the presidential candidates' agenda.

Campaign sweet spot

The business group represents a region along the state's central Atlantic coast that depends heavily on Kennedy, where the space shuttle orbiters are refurbished, mated with solid rockets and a large external tank, and launched. Although the center employs fewer than 2000 civil servants, tens of thousands of locals work for NASA contractors and subcontractors. In addition, shuttle launches draw large numbers of tourists, pumping more money into the local economy.

That prosperity, however, is threatened by the 2004 initiative put forward by President George W. Bush. Under that plan, the shuttle is slated to be retired in 2010 to free up funds for a successor launcher that eventually would take humans back to the moon. On 24 June, NASA Administrator Michael Griffin told a

CREDIT: KYLE CAVALLARO, NASA AMES RESEARCH CENTER



A relaunch. Scientists dedicate the old Moffett Field headquarters as the lunar institute.

New Institute Shoots for the Moon

MOUNTAIN VIEW, CALIFORNIA—The home of the new NASA Lunar Science Institute, which opened its doors here on 1 March, is a shadow of its former glory. Once the centerpiece of the Navy's now-abandoned Moffett Field, the stately stucco building currently sits on the periphery of NASA's Ames Research Center. The two-story structure lacks air conditioning, a conference room, and a working water fountain. But lunar scientists hope to refurbish the shabby surroundings—which now reflect the tattered state of the discipline—as part of a larger renovation that will set the agenda for a new generation of scientific exploration.

Whether that happens depends in large part on the next U.S. president (see main text). The 500 scientists, engineers, and students who gathered here in July to lay out an ambitious new agenda for lunar science are hoping that he retains President George W. Bush's 2004 initiative to return humans to the moon. Lunar scientists hope to ride the coattails of that exploration effort, which will require robots to scout out the lunar environment before astronauts land and conduct extensive research on the surface, beginning in 2020.

But U.S. scientists aren't taking any chances. European, Japanese, and Canadian representatives stayed after the meeting to hammer out plans for an International Lunar Network to coordinate the plethora of lunar missions planned by several nations in the next several years (*Science*, 16 March 2007, p. 1482). So even if the United States ultimately were to bow out of human exploration, its researchers could still have a hand in the field. "During the next 5 years, there will be an astounding amount of data" coming back from the moon, says Carlé Pieters, a planetary scientist at Brown University who co-chaired a National Research Council study last year on lunar research. "There will be a

Senate panel meeting in nearby Port Canaveral that Kennedy alone will lose between 6000 and 7000 shuttle-related jobs. That loss, he added, will be offset by 3000 new jobs at or near the center associated with the space station and the new launcher. An estimated 1000 people converged the night before the hearing for an emotional rally decrying the impact of the shuttle's retirement on local jobs and urging legislators to extend the program.

CREDITS: AP PHOTO/ALEX BRANDON; JOSHUA LOTT/REUTERS/CORBIS; DOMINIC HART, NASA AMES RESEARCH CENTER

"Families are anxious," says Ketcham, who began to woo the campaign staffs of presidential candidates as early as March 2007. "This is not rocket science but simple political arithmetic. This is a critical corridor, and this is an issue which could decide who wins the state." Florida played the decisive role in the tight 2000 race between Bush and Democrat Al Gore and again in 2004 in the race between Bush and John Kerry. Its 27 electoral votes are the fourth biggest prize in the country, a total likely to rise in 2012 after reapportionment following the 2010 census.

Although the demographic trends brought erstwhile Republican candidates Rudy Giuliani and Mitt Romney to the Kennedy Space Center in January, Weatherman knew that simply complaining about job losses was not enough to make space matter in the campaign. "A bigger issue was needed for national leadership to take note," she says. By the time a delegation of Space Coast businesspeople met with McCain staffers in Washington, D.C., in April and 1 month later with Obama staffers in Chicago, Illinois, the commission had found the answer: the projected 5-year gap in access to the international space station between the time the shuttle flies its last mission in 2010 and the new Constellation rocket begins operations by 2015.

To bridge the gap, the White House initiative assumes that U.S. astronauts will hitch rides on the Russian Soyuz vehicles to service the space station. That dependence initially concerned only a handful of politicians, notably U.S. Senator Bill Nelson (D–FL), who flew on the shuttle while a congressman and who is a longtime NASA supporter. Although McCain chaired the Senate Commerce, Science, and Transportation Committee when it held hearings on the proposal in 2004 and 2005, he declined to join the attack on the president's plan. And he remained quiet during the long season of presidential primaries.

But McCain broke his silence in a 29 July statement marking the 50th anniversary of NASA. "My opponent seems content to retreat from American exploration of space for a decade," he declared. "I am not." His reference was to an $18-billion-a-year education plan from Obama that would be paid for in part by delaying the new launcher by an additional 5 years. Although Obama's staff beat a hasty retreat after harsh criticism from Senator Hillary Clinton's (D–NY) staff and space advocates (*Science*, 1 February, p. 565), the candidate himself had not publicly revised his education plan.

To set himself apart, McCain promised to give the Constellation program the funding it needs to begin a new era of human exploration. Although the statement does not

feast—and there are not enough people to analyze it."

With a current budget of $1.5 million, the institute will design neither instruments nor missions, says astronomer David Morrison, its acting director pending the appointment of a permanent head by the end of the year. (NASA has a separate fund for peer-reviewed lunar basic research by individual investigators.) Instead, Morrison says, the institute hopes to be a nexus for a growing number of lunar research teams, complementing other organizations like the Lunar and Planetary Institute in Houston.

In charge. David Morrison is the institute's acting director.

The first visible signs of that commitment will come later this year with the signing of 4-year cooperative agreements with several universities and research institutes. Morrison hopes the institute's 2009 budget will grow to $10 million, split between NASA's science and exploration offices. In addition to funding more data analyses, says Morrison's deputy, Greg Schmidt, the additional resources will help "create a community" of lunar scientists.

That community is eager to provide input for human missions that would explore the moon in far more detail and subtlety than is possible with robotic missions like the current Mars rovers. Scientific questions include the extent and nature of the massive bombardment that took place 3.9 billion years ago, leaving the lunar surface pockmarked; how the lunar crust separated itself from the mantle; and the impact of the ancient solar wind on the lunar surface. Answering such questions requires a human touch. "You can't just send a robot out to collect rocks," says G. Jeffrey Taylor, a planetary scientist at the University of Hawaii, Manoa.

Robotic probes will dominate NASA lunar exploration during the next decade, however. NASA plans to launch the Lunar Reconnaissance Orbiter early next year, a mission that includes Ames's Lunar Crater and Observation and Sensing Satellite. The Gravity Recovery and Interior Laboratory, twin spacecraft designed to map the lunar gravity field in unprecedented detail, will follow in 2011, along with the Lunar Atmosphere and Dust Environment Explorer. Last year, NASA canceled a series of rovers designed to conduct science and provide data on potential human landing sites, but scientists hope to persuade a new Administration to revive them. In the meantime, workers are getting an old building ready for a new mission. **–A.L.**

mention returning humans to the moon, Douglas Holtz-Eakin, a former Congressional Budget Office director and now a senior policy adviser to McCain, said that the Republican "thinks that we need the capacity to put men in space and get to the international space station. He also believes we need to put men back on the moon."

Four days later, Obama came to the Space Coast and picked up the gauntlet that McCain had thrown down. Visiting the nearby town of Titusville with Nelson and standing behind a sign declaring "Economic Security for American Families," the Democratic candidate pledged to allow at least one additional shuttle flight, speed up the shuttle's successor, and make sure "that all those who work in the space industry in Florida do not lose their jobs." He also promised to reestablish a presidential aeronautics and space council "so that we can develop a plan to explore the solar system—a plan that involves both human and robotic missions, and enlists both international partners and the private sector." Standing beside a beaming Nelson, Obama said that "under my watch, NASA will inspire the world, make America stronger, and help grow the economy here in Florida." Nelson and sources close to Obama's campaign say that the candidate wants to boost NASA's budget by $2 billion, although they aren't clear whether that would be a one-time or an annual increase.

McCain didn't wait long to reply. On 12 August, he released a two-page statement in which he promised to finish the space station, support Constellation, and ensure that space exploration "is a top priority." He also touted his history of pressing NASA to control costs and promised to prevent pork-barrel spending from sapping the agency's research muscle. And on 18 August, McCain spent an hour in a closed-door session with space industry representatives adjacent to Kennedy, where Weatherman says he questioned them on NASA's future. At a press conference the same day, he criticized Obama for having a "short, thin record" on space.

By then, however, Obama had released a seven-page paper laying out in surprising detail his plan for space. Along with backing a new generation of science probes and observatories, the plan endorsed sending humans to the moon by 2020—as Bush proposed—and eventually on "to more distant destinations, including Mars." To get there, Obama backed a new launcher, though he did not endorse the specific Constellation effort, which faces technical and budgetary hurdles. Two weeks later, he repeated that message in a reply to questions from a grassroots science advocacy campaign, calling it part of "a 21st century vision of space" (Sciencedebate2008.com).

McCain has also continued to speak out. Two days after his visit to the Space Coast, he told a Florida television station that "I stand for not cutting any of the NASA budget, which Senator Obama proposed and then reversed himself, as he has on a number of things." On the same day, Obama explained that he changed his view after consultations.

Heavy lifting. The potential loss of thousands of jobs at Kennedy Space Center in Florida, where workers prepare the shuttle orbiter Atlantis for its next launch, has helped Lynda Weatherman (inset) and other local business leaders inject space into the presidential campaign.

Making the case

Obama's attention to detail, say several Democratic insiders, owes much to a small group of space advocates led by Lori Garver, a former NASA official under President Bill Clinton, director of the National Space Society, and currently a consultant with Washington's Avascent Group, an organization that works primarily for aerospace companies. Garver, who advised Senator Clinton's campaign until her defeat this spring, declined to discuss how that policy was formulated. Her counterpart on the McCain campaign, former Apollo astronaut Walter Cunningham, isn't so shy. He calls Garver "the architect" of Obama's space policy, which he said shows "a good understanding of space science and exploration."

Meanwhile, the conflict between Russia and Georgia, which broke out in early August, provided an unexpected boost to those hoping to make the launch gap a campaign issue. The resulting deterioration in U.S.-Russian relations led a growing number of politicians to question NASA's dependence on the Soyuz. McCain and two other senators sent Bush a 25 August letter warning that retiring the shuttle promptly could endanger U.S. access to the space station even if it makes financial sense. Speeding up the new launcher program, which already faces technical hurdles, or encouraging private launchers won't be enough to close the gap, they warned. Instead, they recommended that NASA "take no action for at least one year from now that would preclude the extended use of the space shuttle beyond 2010." The message, says Cunningham, is "we gotta keep the shuttle flying."

Griffin complained bitterly in a recent e-mail that White House officials are conducting a "jihad" against the shuttle and that the only "politically tenable course" for the next president is to extend the shuttle, according to a report in the 7 September *Orlando Sentinel*. Garver admits that she is "thoroughly amazed" the gap has emerged as a contentious national issue.

Unlike in past campaigns, space has even become part of each party's platform, a nonbinding compilation of positions. There's a brief mention in the Democratic version of a "strong and inspirational vision" for space and a sentence in the Republican document declaring that "we look toward our country's return to the moon." That's nice, say the Florida Space Coasters, but it's not enough. "[The candidates] are competing with one another, which is good for us," says Ketcham. Adds Weatherman, "Now we need specifics."

–ANDREW LAWLER

CREDITS: NASA/JACK PFALLER; MICHAEL G. WOOLLEY/UNITED LAUNCH ALLIANCE

Second nature. Reared at NARC and set free in the UAE desert, this houbara bred 2 years later.

CREDIT: NARC

ORNITHOLOGY

The Houbara: Headed for Oblivion?

The elusive Asian houbara bustard could fall victim to falconers and poaching without strong international protection

URUMQI, CHINA—When Yang Weikang stalks his quarry in the Junggar Basin of western China, he needs all the patience he can muster. "The creature is shy—and *very* cunning," says Yang, an ecologist at the Xinjiang Institute of Ecology and Geography of the Chinese Academy of Sciences in Urumqi. The elusive animal is the Asian houbara bustard (*Chlamydotis undulata macqueenii*), a cranelike bird with sandy buff plumage, mottled with dark-brown spots, that nests in open desert and dry steppe. Yang's team uses telescopes for observations; with its superb vision and a clear line of sight, houbaras can spot threats from hundreds of meters away.

But the houbara's guile alone will not save it from oblivion. The bird has the unhappy fate of being the favorite prey of falconers. Over the past few decades, hunting pressure across a wide swath of Asia has risen in concert with two other threats: poaching and habitat loss as arid land is converted to farms or urban sprawl.

To address these woes, bird experts are negotiating with governments to establish protected areas in key countries where the bird breeds or winters. "We're working on this very seriously," says behavioral ecologist Olivier Combreau, director of the National Avian Research Center (NARC) in Abu Dhabi, United Arab Emirates (UAE). Creating new reserves where taking houbaras is banned and enacting stiffer penalties for poaching and overhunting are components of an action plan the signatories of the Convention on the Conservation of Migratory Species of Wild Animals are now reviewing. "If we do nothing, there is no hope for the houbara," says Yang.

The houbara's downward spiral began with the economic rise of the Persian Gulf, and it accelerated after the Cold War ended. Adults, which are about 60 cm long, have a wingspan of 140 cm, and weigh in at around 2 kg, breed in early spring in China, Kazakhstan, and Mongolia. Around late September, houbaras head south on a journey of up to 7000 km. Some migrate to Iran and the Arabian Peninsula, while others flock to Afghanistan and Pakistan.

Houbaras that winter in the Persian Gulf often end up on dinner plates. Falcons are trained to hunt the delicacy. Falconry, an Arab tradition, soared in popularity as the oil-producing nations grew rich; falconers, who until the 1960s struck out on horse or camel, now roll into the desert in four-wheel drive convoys. "Before oil, hunting was a way to make a living," says Yang. "Now it's mostly for sport." As houbaras became scarcer, falconers descended on wintering grounds in Pakistan. The Soviet disintegration in 1991 opened a new frontier; newly independent central Asian countries were soon welcoming hunting parties to bag houbaras during its spring and fall migrations.

Researchers knew the embattled bird was on the ropes. But when a team led by Combreau undertook surveys in China, Kazakhstan, and Oman from 1998 to 2002, they discovered that houbaras were vanishing before their eyes. During the 4-year study, houbara numbers declined by 63% in China, 60% in Kazakhstan, and 50% in Oman, the researchers reported in *Biological Conservation* in 2005. Some experts have pegged the Asian population at about 50,000, but Combreau says no one really knows how many are left. One thing is certain, he and his colleagues warned in their 2005 paper: "The Asian houbara may face extinction in the wild in the foreseeable future."

The houbara's decline has continued over the past few years, Combreau says. His team keeps a close eye on Kazakhstan, where their surveys show drops of between 5.5% and 8.3% a year. Because most houbaras breed in Kazakhstan and the rest migrate through the vast country, decreases there "give a fairly good idea" of the overall population's vulnerability, Combreau says. As a last resort, he says, NARC's successful captive-breeding program could reintroduce houbaras into areas in the wild where the bird goes locally extinct.

Averting that doom will mean reining in falconers. In China and Mongolia, taking houbaras is outlawed. In UAE, hunting is limited to a couple of months a year in a few spots. "They will not stop the hunt, of course. It's tradition," Combreau says. "But there is a genuine effort here to promote sustainable hunting." Many other countries regulate hunting, he says, but lack the means for enforcement.

One bright spot is Pakistan. Falconers are a rare sight there these days because of the deteriorating security situation. "Last year, hardly any hunting took place," says Mukhtar Ahmed, president of Houbara Foundation International Pakistan in Lahore. And poaching has declined, he says, because Gulf nations have cracked down on the market for houbaras used to train falcons. Still, NARC estimates that up to 7000 houbaras each year are spirited into UAE alone.

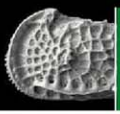
The houbara's decline pains Yang, who holds the bird in high esteem. To lure a fox or other predator away from her nest, a female will bravely hobble away from her eggs, pretending to have a broken wing. Once the predator is off the scent, she'll drop the charade and run back to the nest. A female once put similar moves on Yang's jeep: "She ran slowly in front of us, trying to guide our car away from her nest," he says.

Such rare encounters require spending weeks in the field. In China and Kazakhstan, there's roughly one bird per 10 square km. "It's very unlikely one would see a houbara in the wild," Combreau says, except when males are putting on breeding displays for females. Otherwise, "they're almost impossible to detect." Unless countries act quickly and forcefully, it soon may be impossible even for the falcons to detect the furtive, and fading, houbaras.

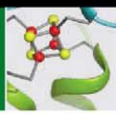
–RICHARD STONE

Addressing Earth's history 1447

Understanding body size 1451



Reduced iron-sulfur clusters 1452



LETTERS | BOOKS | POLICY FORUM | EDUCATION FORUM | PERSPECTIVES

LETTERS

edited by Jennifer Sills

Working the Crowd

THE NEWS FOCUS STORY BY J. TRAVIS ("SCIENCE BY THE MASSES" 28 March, p. 1750) describes the application of crowd-sourcing in research and development. This approach consists of gathering a mass of people to seek out new ideas or solutions, and paying profits to seekers and solvers. A related idea is crowd-funding, a bottom-up model of financing used for various purposes, from software development to political campaigns.

We suggest crowd-funding as a possible strategy to cope with the lack of investments in research, as well as to increase democratization in the sciences. Projects seeking funding could be stored in an online repository. Each project would include a description of its objectives, duration, and requested contribution. Investors (either people or funding agencies) could decide which projects to fund.

For such a service to be successful, several challenges would need to be addressed: (i) Evaluating the quality of the proposals. To assist (nonspecialist) investors in deciding the awarding of contributions (and to audit thereafter), a peer-review procedure could be used. (ii) Potential for fraud. Fraud could be prevented by implementing a reputation system (*1*) and by indicating the scientific track record of the proponent. (iii) Intellectual property management. Intellectual property issues could be managed by allowing proponents to choose the appropriate level of protection of their ideas—for example, by using Creative Commons licenses (*2*). (iv) Investor rewards. Investors could be motivated by the prospect of earning shares (for profit-making research programs) or by the acknowledgment of their contribution (for nonprofit research programs).

Public investing. Allowing the public to invest in research may help alleviate the funding shortage.

ANDREA GAGGIOLI AND GIUSEPPE RIVA

Department of Psychology, Catholic University of Milan, Milan 20123, Italy, and Applied Technology for Neuro-Psychology, Istituto Auxologico Italiano, Milan 20145, Italy.

References

1. P. Resnick, R. Zeckhauser, E. Friedman, K. Kuwabara, *Comm. ACM* **43**, 45 (2000).
2. Creative Commons (www.creativecommons.com).

Southern Ocean Not So Pristine

THE REPORT "A GLOBAL MAP OF HUMAN impact on marine ecosystems" (B. S. Halpern *et al.*, 15 February, p. 948) provides a timely overview of anthropogenic effects on even the farthest reaches of Earth's oceans. However, we contend that, for at least one region, using data from only the past decade leads to misleading results.

A widespread perception exists that waters south of the Antarctic Polar Frontal Zone—i.e., the Southern Ocean (SO)—are still nearly pristine (*1*, *2*). In fact, the northern portion of the SO saw virtually all cetacean populations removed long ago (*3*), and in subsequent years (1960s to 1980s) the largest stocks of demersal fish in the Indian Ocean and Scotia Sea/Atlantic Ocean sectors were also fished to commercial extinction (*4*, *5*). Historically exploited fish species and cetaceans show little signs of recovery in the SO, and recent legal commercial fishing activity has been correspondingly low (*6*). It is thus no surprise that the modeling used by Halpern *et al.* shows little anthropogenic impact in these sectors apart from that of climate change. The authors acknowledge that accounting for current illegal, unregulated, and unreported fishing in these waters might show increased human impacts. The additional consideration of historical data should cause Halpern *et al.* to temper their conclusion that for the world's oceans "large areas of relatively little human impact remain, particularly near the poles."

LOUISE K. BLIGHT[1] AND DAVID G. AINLEY[2]

[1]Centre for Applied Conservation Research, University of British Columbia, Vancouver, BC V6T 1Z4, Canada. [2]H. T. Harvey and Associates, Los Gatos, CA 95032, USA.

References

1. J. P. Croxall, P. N. Trathan, E. J. Murphy, *Science* **297**, 1510 (2002).
2. V. Smetacek, S. Nicol, *Nature* **437**, 367 (2005).
3. L. Ballance *et al.*, in *Whales, Whaling and Ocean Ecosystems*, J. A. Estes, D. P. Demaster, D. F. Doak, T. E. Williams, R. L. Brownell Jr., Eds. (Univ. of California Press, Berkeley, CA, 2006), pp. 215–230.
4. O. Gon, P. C. Heemstra, *Fishes of the Southern Ocean* (J. L. B. Smith Institute of Ichthyology, Grahamstown, South Africa, 1990).
5. K.-H. Kock, *Antarctic Fish and Fisheries* (Cambridge Univ. Press, Cambridge, 1992).
6. Fishery Reports, *Convention on the Conservation of Antarctic Marine Living Resources* (www.ccamlr.org/pu/e/e_pubs/fr/drt.htm).

Diminishing Sea Ice

IN THEIR USEFUL REPORT, "A GLOBAL MAP OF human impact on marine ecosystems" (15 February, p. 948), B. S. Halpern *et al.* wrote that "large areas of relatively little human impact remain, particularly near the poles." They failed to take into account sea-ice diminishment, which may already be responsible for substantial local, regional, and global effects (*1*, *2*).

Arctic and Antarctic sea-ice ecosystems, together covering 7% of Earth, comprise "one of the largest biomes on Earth" (*3*), providing

CREDIT: JUPITER IMAGES

habitat for many species, from epontic algae to ice-dependent pinnipeds. Recent sea-ice diminishment has been a consequence, in part, of human greenhouse-gas production, indicating that sea ice may also be viewed as an anthropogenic driver of change (*4*, *5*), which will predictably have cumulative and synergistic effects on shelf seas, neighboring ecosystems, and regional to global climate.

Over Beringia (the combined shelves of the Bering and Chukchi seas), diminishment of sea ice may have already reached a "tipping point" (*6*). Multiple effects are apparent. Sea ice provides breeding, feeding, and molting habitat for polar pinnipeds; ribbon seals and Pacific walruses are being considered for threatened or endangered status under the Endangered Species Act. Polar bears have already been designated as "threatened." Ecosystem effects, such as diminished productivity (*7*) and loss of walrus mixing of benthic sediment important to its structure and chemical exchanges (*8*), are also probable. Socioeconomic effects include the relocation of Alaskan coastal villages due to shore erosion, and losses of critical resources on which indigenous subsistence hunters depend, with cascading impacts on hunting practices, knowledge systems, and cultures (*9*, *10*). Commercial species are also impacted by a shift from an Arctic to a sub-Arctic ecosystem (*11*).

Sea ice should not be omitted from consideration in such efforts as this mapping effort exemplifies.

G. CARLETON RAY,[1] GARY L. HUFFORD,[2] IGOR I. KRUPNIK,[3] JAMES E. OVERLAND[4]

[1]Department of Environmental Sciences, University of Virginia, Charlottesville, VA 22904, USA. [2]National Weather Service, NOAA, Anchorage, AK 99513, USA. [3]Arctic Studies Center, Smithsonian Institution, Washington, DC 20560, USA. [4]NOAA Pacific Marine Environmental Laboratory, Seattle, WA 98115, USA.

References

1. *Arctic Climate Impact Assessment* (Cambridge Univ. Press, Cambridge, 2005).
2. J. E. Overland, M. Wang, *Geophys. Res. Lett.* **34**, L17705 (2007).
3. D. N. Thomas, G. S. Dieckmann, *Sea Ice: An Introduction to Its Physics, Chemistry, Biology and Geology* (Blackwell Publishing, Oxford, 2003).
4. W. L. Chapman, J. E. Walsh, *J. Clim.* **20**, 609 (2007).
5. J. E. Overland, M. Wang, S. Salo, *Tellus*, 10.1111/j.1600-0870.2008.00327 (2008).
6. R. W. Lindsay, J. Zhang, *J. Clim.* **18**, 4879 (2005).
7. J. M. Grebmeier *et al.*, *Science* **311**, 1461 (2006).
8. G. C. Ray, J. McCormick-Ray, P. Berg, H. E. Epstein, *J. Exp. Mar. Biol. Ecol.* **330**, 403 (2006).
9. I. Krupnik, in *Impact of Changes in Sea Ice and Other Environmental Parameters in the Arctic* (Marine Mammal Commission, Washington, DC, 2000), pp. 25–39.
10. I. I. Krupnik, D. Jolly, *The Earth Is Faster Now: Indigenous Observations of Arctic Environmental Change* (ARCUS, Fairbanks, AK, 2002).
11. F. J. Mueter, M. A. Litzow, *Ecol. Appl.* **18**, 309 (2008).

Response

A KEY MOTIVATION FOR OUR RESEARCH WAS to counteract the tendency to focus on single activities or single ecosystems when assessing the state of the oceans. Because we mapped cumulative impacts of 17 different stressors on 20 ecosystems, cumulative impacts diverge substantially from expectations for single stressors or ecosystems in most locations. Our global map is particularly valuable because it allows different areas of the planet to be compared using the same currency. The poles, while not pristine, are areas of little cumulative impact relative to the rest of the world's oceans according to our model.

Blight and Ainley suggest that our maps for the Southern Ocean are misleading because they do not reflect current or historical levels of fishing. Had data for historical fishing been available globally and included, many key areas of the world's oceans (e.g.,

waters around Europe, Asia, and North America) would likely have also looked much worse on our map, keeping the patches of blue near Antarctica in the lowest category of human impact. Moreover, the nine stressors not related to commercial fishing and climate change are inarguably lower in the poles than elsewhere. Consequently, our results are conservative, not misleading.

Ray *et al.* highlight a key challenge in mapping human impacts to marine ecosystems at the global scale: capturing dynamic processes on a static map. Sea ice is clearly an important ecosystem, but its extent shows strong seasonal and annual variation, far more than any other ecosystem, precluding a clear solution for where to place it on the map. Ray *et al.* also point out that the loss of sea ice is an important driver of change, but habitat loss and resulting species extirpations are problems shared by many ecosystem types (e.g., mangroves). Future efforts to refine our maps would benefit from incorporation of dynamic processes and may be able to better include historical data when conducted at local or regional scales.

BENJAMIN S. HALPERN,[1]* CARRIE V. KAPPEL,[1] FIORENZA MICHELI,[2] KIMBERLY A. SELKOE,[1,3] CATERINA D'AGROSA,[4] JOHN BRUNO,[5] KENNETH S. CASEY,[6] COLIN M. EBERT,[1] HELEN E. FOX,[7] ROD FUJITA,[8] DENNIS HEINEMANN,[9] HUNTER S. LENIHAN,[10] ELIZABETH M. P. MADIN,[11] MATT PERRY,[1] ELIZABETH R. SELIG,[4,12] MARK SPALDING,[13] ROBERT STENECK,[14] SHAUN WALBRIDGE,[1] REG WATSON[15]

[1]National Center for Ecological Analysis and Synthesis, Santa Barbara, CA 93101, USA. [2]Hopkins Marine Station, Stanford University, Pacific Grove, CA 93950–3094, USA. [3]Hawai'i Institute of Marine Biology, Kane'ohe, HI 96744, USA. [4]School of Life Sciences, Arizona State University, Tempe, AZ 85287–4501, USA. [5]Department of Marine Sciences, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599–3300, USA. [6]National Oceanographic Data Center, National Oceanic and Atmospheric Administration (NOAA), Silver Spring, MD 20910, USA. [7]Conservation Science Program, World Wildlife Fund–United States, Washington, DC 20037, USA. [8]Environmental Defense, Oakland, CA 94618, USA. [9]Ocean Conservancy, Washington, DC 20006, USA. [10]Bren School of Environmental Science and Management, University of California, Santa Barbara, CA 93106, USA. [11]Department of Ecology, Evolution, and Marine Biology, University of California, Santa Barbara, CA 93106, USA. [12]Curriculum in Ecology, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599–3275, USA. [13]Conservation Strategies Division, The Nature Conservancy, Newmarket CB8 8AW, UK. [14]School of Marine Sciences, University of Maine, Darling Marine Center, Walpole, ME 04353, USA. [15]Fisheries Center, University of British Columbia, Vancouver, BC V6T 1Z4, Canada.

*To whom correspondence should be addressed. E-mail: halpern@nceas.ucsb.edu

Microscopy for Life Scientists

IN THE REPORT "HIGH-RESOLUTION SCANning x-ray diffraction microscopy" (18 July, p. 379), P. Thibault *et al.* described the technique of scanning x-ray diffraction microscopy (SXDM) and their achievement of 70-nm resolution. The implications of this work for life scientists were underemphasized.

The development of SXDM comes from the physics community, but it has the most to offer to the life science community, where the inability to image live cells and tissues beyond the limitations of the light microscope is a constant frustration. If SXDM achieves its potential resolution of below 10 nanometers on live material, it will result in a revolution. In addition to refining their technique, I hope Thibault *et al.* will start to apply their existing technique to live biological tissues with inherent periodicity (such as muscle or cornea), where it already has the potential to provide unique information.

NIGEL J. FULLWOOD

Department of Biomedical Sciences, Lancaster University, Lancaster LA1 4YQ, UK. E-mail: n.fullwood@lancaster.ac.uk

Archaeology Without Borders

I READ WITH INTEREST THE LATEST UPDATES, problems, and progress in the archaeological research in the neighboring countries of India and Pakistan (Special News Focus section, "Unmasking the Indus," 6 June, p. 1276). However, I was saddened to note the researchers' apathy about the international border that prevents the free flow of information between sides. Sites in this area are not just important to India and Pakistan; these are precious global heritage centers marking the triumph of human civilization and evolution of human ingenuity and progress.

The international community should come together and work as a consortium to facilitate easy access to information for all who research these sites, regardless of nationality or location. It is important for all of us to understand our shared heritage and to preserve it for future generations to learn and appreciate.

SAIKAT KUMAR BASU

Department of Biological Sciences, University of Lethbridge, Lethbridge, AB T1K 3M4, Canada. E-mail: saikat.basu@uleth.ca

Letters to the Editor

Letters (~300 words) discuss material published in *Science* in the previous 3 months or issues of general interest. They can be submitted through the Web (www.submit2science.org) or by regular mail (1200 New York Ave., NW, Washington, DC 20005, USA). Letters are not acknowledged upon receipt, nor are authors generally consulted before publication. Whether published in full or in part, letters are subject to editing for clarity and space.

TECHNICAL COMMENT ABSTRACTS

COMMENT ON "A Global Map of Human Impact on Marine Ecosystems"

Michael R. Heath

Halpern *et al.* (Reports, 15 February 2008, p. 948) integrated spatial data on 17 drivers of change in the oceans to map the global distribution of human impact. Although fishery catches are a dominant driver, the data reflect activity while impacts occur at different space and time scales. Failure to account for this spatial disconnection could lead to potentially misleading conclusions.

Full text at www.sciencemag.org/cgi/content/full/321/5895/1446b

RESPONSE TO COMMENT ON "A Global Map of Human Impact on Marine Ecosystems"

Kimberly A. Selkoe, Carrie V. Kappel, Benjamin S. Halpern, Fiorenza Micheli, Caterina D'Agrosa, John Bruno, Kenneth S. Casey, Colin Ebert, Helen E. Fox, Rod Fujita, Dennis Heinemann, Hunter S. Lenihan, Elizabeth M. P. Madin, Matt Perry, Elizabeth R. Selig, Mark Spalding, Robert Steneck, Shaun Walbridge, Reg Watson

Our results provide an important first step toward a full assessment of how human activities act cumulatively to affect the condition of the oceans. Fisheries (and climate change) impacts are some of the hardest to map and measure accurately. Consequently, species-specific considerations and fine-scale analyses should be left to more nuanced regional-scale replicates of our mapping framework.

Full text at www.sciencemag.org/cgi/content/full/321/5895/1446c

CORRECTIONS AND CLARIFICATIONS

Brevia: "Auxin gradients are associated with polarity changes in trees," by E. M. Kramer *et al.* (20 June, p. 1610). E. M. Kramer should have been affiliated with both Bard College at Simon's Rock, Massachusetts, and the Centre for Plant Integrative Biology at the University of Nottingham, UK.

HISTORY OF SCIENCE

Illuminating the Details of Deep Time

Ralph J. O'Connor

Among the tools required by the paleontologist, probably not many people today would include pastry. In January 1828, in the well-appointed home of Roderick Murchison, the enterprising geologist William Buckland demonstrated to the assembled savants that fossilized footprints recently discovered in sandstone near Dumfries belonged to ancient tortoises. To do this, he rolled out a large sheet of "soft pye-crust," produced three tortoises and a reluctant crocodile, and invited them to walk across the crust. After some prodding, the animals set off, only to get stuck in the pastry. The cooks got back to work, as one guest recalled:

> It was really a glorious sight to behold all the philosophers, flour-besmeared, working away with tucked-up sleeves. Their exertions, I am happy to say, were at length crowned with success; a proper consistency of paste was attained, and the animals walked over the course in a very satisfactory manner; insomuch that many who came to scoff returned rather better disposed towards believing.

This was no mere publicity stunt. Buckland's experiment was an effective application of the geological principle of actualism: using evidence and analogies from the present to unlock the past. In itself it is no sillier, and a good deal less far-fetched, than more recent attempts to replicate the origin of life by synthesizing a "primeval atmosphere" and setting off spark plugs in it.

Archibald Geikie's time-honored mantra "the present is the key to the past" could serve as a motto for Martin Rudwick's *Worlds Before Adam*. In his previous book, the critically acclaimed *Bursting the Limits of Time* (*1*), Rudwick explained how European savants in the late 18th and early 19th centuries (the "Age of Revolution") created geology as a historical science, borrowing tools from disciplines such as biblical chronology and civil history to shed new light on rocks and fossils. Thus was born the concept of "former worlds" inhabited by extinct creatures, with French comparative anatomist Georges Cuvier starring as chief midwife. That book ended with Buckland's celebrated identification of a fossil hyenas' den in a Yorkshire bone cave: this snapshot of a former world in all its gory detail was widely felt to realize Cuvier's prophecy that geologists would one day rival Newton's achievement in space and "burst the limits of time."

Worlds Before Adam
The Reconstruction of Geohistory in the Age of Reform

by Martin J. S. Rudwick

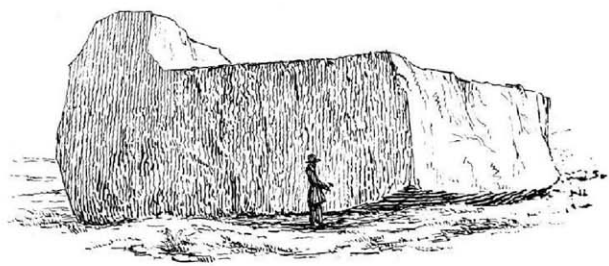
University of Chicago Press, Chicago, 2008. 638 pp. $49, £25.50. ISBN 9780226731285.

In the sequel, Rudwick traces the new science's rise to maturity. He shows how Cuvier's successors in the period 1820–1845 put their new tools into practice with increasing sophistication, illuminating the depths of a still more remote antiquity. As the evidence piled up, it became increasingly apparent that there was a progressive trend in the history of life on Earth, with more and more complex forms emerging over time—although the most prominent figure in Rudwick's story, Charles Lyell, resisted these conclusions for a surprisingly long time, in part because he was so worried about their possible evolutionary connotations.

Despite Lyell's anxieties, the grand creation story thus revealed was perfectly compatible with Christian understandings of God creating species over unfathomed aeons rather than six literal days. Deep time had no angst attached: the leading geologists, who included several churchmen, unanimously accepted the vast scale of geohistory. (Some other churchmen were less enamoured of the new geology, but that is another story.) Most of Rudwick's key players were religious believers, and many adhered to fairly conservative varieties of Christianity. But far from retarding or compromising their science, their religious backgrounds had enabled them to make the imaginative leap needed to see their planet as an object in history, with a past and a future.

Their search for traces of Noah's Flood may seem laughable to a present-day geologist, but Rudwick convincingly shows how the physical evidence available in the 1820s really did point toward a massive, unprecedented cataclysm of some kind. The riddle was solved in the 1840s, when the Swiss (and Calvinist) savant Louis Agassiz introduced the idea of a global Ice Age into mainstream geological debate. The Flood froze over and "diluvial currents" morphed into glaciers, but the scientific methods remained the same. The geohistory hammered out by these savants has come down to us unchanged in its essentials, although many more details have since been filled in.

Like its predecessor, *Worlds Before Adam* is the product of painstaking research. It appears dauntingly long but is a delight to read. Rudwick's style is lucid and engaging throughout, and he is unfailingly courteous to his nonspecialist readers, ensuring that all terms and concepts are fully explained and avoiding unnecessary jargon. The book's strictly chronological arrangement gives it a strong narrative thrust, and its many beautifully printed illustrations and generous quotations from original sources enhance the sense of primary contact with the evidence.



CREDIT: H. T. DE LA BECHE (2)

Carried on a flood? In the early 19th century, the best explanation for the innumerable erratics found many miles from their sources appeared to be violent aqueous events that carried masses of floating ice charged with boulders. This huge block of granite above Lake Como, Italy, was sketched by Henry Thomas De la Beche (*2*).

The reviewer is at the Department of History, University of Aberdeen, Crombie Annex, Meston Walk, King's College, Old Aberdeen AB24 3FX, UK. E-mail: ralph.j.oconnor@gmail.com

Rudwick tells us that his aim was not to produce a definitive history of geology as a whole, but rather to explain how geologists pieced together a long history for the Earth we inhabit. This is no mean feat in itself. He makes a good case for seeing the "Cuvier-Lyellian revolution" as a paradigm shift no less important than those associated with Copernicus, Darwin, and Freud, but one which has attracted far less attention—perhaps because it has so often been treated as a mere backdrop for the rise of evolutionary theory. In these two graceful and judicious volumes, the culmination of a distinguished career, Rudwick has restored geology to its rightful historical place at the heart of modern scientific culture. More than this, he enables readers to experience geology as a new science. By immersing us in the investigations, reflections, and debates of the time, he lifts us out of our present-day perspective so that we see the objects of geology afresh, through the astonished eyes of those who created it.

References

1. M. J. S. Rudwick, *Bursting the Limits of Time: The Reconstruction of Geohistory in the Age of Revolution* (Univ. Chicago Press, Chicago, 2005); reviewed by N. Oreskes, *Science* **314**, 596 (2006).
2. H. T. De la Beche, *Sections and Views, Illustrative of Geological Phaenomena* (London, 1830).

10.1126/science.1162473

SCIENCE COMMUNITY

Scientists Under Siege

Deborah C. Runkle

Suppose you are a scientist and a finalist for the position of vice president for research at the University of South Florida (USF). Before leaving for your interview trip, you receive copies of letters sent to the university's administration informing them of your "ignominy" and stating that you are unwelcome in the university's town. Animal rights activists meet your plane and (because of an open meetings law) are present at most of your interviews. Activists outside the meeting room doors lobby attendees and distribute fliers that make false and preposterous claims about your research. Demonstrators wear T-shirts demanding that you not be hired. When you deny the accusations being hurled at you, a faculty member calls you a "son of a bitch" and a liar. At your hotel room, you receive threatening calls and knocks on your door in the middle of the night. Fortunately, the campus police provide you with protection. Arriving at the airport for your return trip, you are surrounded and harassed by demonstrators until airport security rescues you. At home, you find protesters standing not far from your house, shouting at you. And USF's president now refuses to speak to you. You don't get the job.

All this and more happened to P. Michael Conn, an author of *The Animal Research War*. Conn (associate director of the Oregon National Primate Research Center and a professor at Oregon Health and Science University) used to conduct research using animals, but has been studying cell lines for many years. His co-author, James Parker, was formerly the primate center's public affairs officer.

The Animal Research War

by P. Michael Conn and James V. Parker

Palgrave Macmillan, New York, 2008. 219 pp. $34.95, £22.99. ISBN 9780230600140.

Their important book could not have appeared at a more auspicious time. In the past, animal rights extremists primarily targeted industry and academic research facilities. But as these institutions have become better at protecting themselves from physical attacks, the extremists' crosshairs are now squarely on individual scientists.

Nowhere is this more true than in California, where protesters have conducted "home visits": trespassing, parading in front of houses while shouting that the scientists are torturers and murderers, and distributing leaflets to researchers' neighbors that make the same accusations. Scientists and their families have received threatening e-mails and phone calls. Cars and homes of researchers at the University of California, Santa Cruz, have recently been fire bombed, forcing one neurobiologist and his family to escape their smoke-filled house (*1*, *2*). Previously, one University of California, Los Angeles, primate neurobiologist had announced that he would no longer conduct research on animals, asking the extremists to stop making threats against his children (*3*).

Conn and Parker chronicle the escalating attacks on scientists and their families. The book, however, is more than just a list of horrors. The authors provide a concise yet comprehensive review of issues relating to animal research and the opposition to it. They spell out the history and strategies of the animal welfare and animal rights movements—and the difference between these. They offer biographical sketches of their dramatis personae and summarize the philosophical arguments for and against animal rights and "personhood" for animals.

The authors explain what basic research is and why it sometimes involves the use of living animals. They describe important biomedical advances that have relied on animals along with the "startling audacity" by which animal rightists "reinvent the great stories of biomedical triumphs." They discuss the laws, regulations, and inspection regimes governing the use of animals in the United States, as well as the voluntary policing mechanisms in which nearly all industrial and academic research institutions participate. (This section would have benefited from the inclusion of more details, boring as those sometimes are.)

Importantly, Conn and Parker take universities and professional societies to task for not playing an active and positive role in informing the public about animal research. That said, they have an overly optimistic view of how more information about such research will change the "climate of fear." The public no more wants to know the details about how their medicines made it to the pharmacy than they want to know how their roasting chicken made it to the grocery store. Missing from this otherwise excellent account is an admission that some researchers and their institutions have fallen short of the high standards set by their professional responsibilities and by the government. These situations do occur, albeit in a minority of cases, and a discussion of the issue would have increased the authors' credibility.

The book is well researched and documented. One appendix answers 20 frequently asked questions about animal research—e.g., "How can research results derived from animal research be applied to humans?" "Aren't lost and stolen pets used in research?"—and a second provides Web addresses of relevant groups (on both sides of the argument) for those wanting to know more. The authors' prose is sometimes dramatic, sometimes caustic, sometimes humorous, only occasionally treacly, and always sharp and snappy. *The Animal Research War* offers an invaluable resource that will not be soon replaced.

References

1. G. Miller, *Science* **321**, 755 (2008).
2. M. R. C. Greenwood, G. Ringold, D. Kellogg, *Science* **321**, 891 (2008).
3. *Science* **313**, 913 (2006).

10.1126/science.1165110

The reviewer is with the Science and Policy Programs, American Association for the Advancement of Science, 1200 New York Avenue NW, Washington, DC 20005, USA. E-mail: drunkle@aaas.org

ETHICS

Do We Need "Synthetic Bioethics"?

As we address ethical issues in emerging fields, ethicists, funders, and policy-makers should resist balkanization.

Erik Parens,* Josephine Johnston, Jacob Moses

With the explosion of public interest in human genetics that surrounded the launch of the Human Genome Project, a new field of bioethics was born, and named "gen-ethics." Shortly after the end of the Decade of the Brain in the early 2000s, neuroethics was born. Soon after came nano-ethics. Now, synthetic biology is the hot new star, and calls for examination of its ethical implications are growing louder.

In the wake of the announcement earlier this year of synthesis of a bacterial genome, two German scholars suggested that synthetic biology has ethical implications distinct from those raised by genetic engineering (*1*). A British group has just published a White Paper on synthetic biology's social and ethical challenges (*2*). Researchers funded by the European Commission recently hosted an electronic conference devoted to safety, security, and ethical concerns associated with synthetic biology (*3*). The 2006 and 2007 International Meetings on Synthetic Biology featured presentations on ethical issues, and the 2008 conference is scheduled to include sessions on security, societal issues, and policy. The Hastings Center has recently received two grants to map the ethical issues in synthetic biology (*4*), including one from the Alfred P. Sloan Foundation, which is considering a larger initiative in this area.

Is it time for the birth of yet another bioethical subfield, perhaps "synthetic bio-ethics"? Although creating such a subfield might be in the short-term self-interest of bioethicists, in the long run, further balkanization of bioethics would be a mistake.

Asking bioethical questions in the context of emerging science and technology is hugely important for our health, environment, and, ultimately, our democracy. But anyone who engages with those questions must acknowledge the extent to which they are similar from one scientific arena to another. After all, if synthetic biologists are able to create biofactories that make gene products, they are engaging in a form of genetic engineering that, presumably, could be considered in gen-ethics. Insofar as synthetic biologists work at the nanoscale, their work seems to fall within the purview of nano-ethics, and so on. Given the convergence of scientific investigations, it is not logical to separate the associated ethical inquiries.

Particular ethical questions are certainly more pressing in some arenas than in others. For example, concerns about privacy might be more pressing in genetics than in synthetic biology and concerns about civil liberties might be more pressing in neuroscience than in genetics. But the questions themselves are virtually identical to the ethical questions that have arisen in the past. Failing to recognize that fact can lead to reinventing the bioethical wheel for each new technology and, thus, squandering scarce resources. Instead of lovingly listing the ethical questions that arise over and over, we need to dig deeper. We need to test intuitions, arguments, and responses developed in previous contexts against new fact patterns.

When bioethicists think they have found a new set of ethical questions, they are prone to think they can provide a new set of answers. We are not here wagging our fingers at others; one of the authors of this piece has himself, in once promising new guidelines for prenatal genetic testing, fallen prey to such irrational exuberance (*5*). If we better recognized that the questions are, at core, familiar, we might make more realistic promises, better appreciating salient differences among the kinds of responses we can offer.

Some of the ethical questions raised by synthetic biology are of a kind that we have addressed with some success. Whether safety should be left to self-regulation (*6–9*) or conducted by a public body is a familiar ethical question. And the good news is that we have experience to draw on. For example, we have learned from tobacco, asbestos, and pharmaceutical products that information about risk is not always shared voluntarily. Our concern to protect people from harm and preserve informed consumer choice can legitimately outweigh our commitment to minimal regulation and free markets, which is why we sometimes compel sharing of risk information. In genetics, we have learned that individuals can react to the same risk information in fundamentally different ways. With nanotechnology, we are learning that there is a public expectation and a private-sector desire for international standards and rigorous risk assessment. Focus groups have shown public demand for independent, third-party risk assessment of emerging technologies, which calls into question claims that self-regulation will suffice to reassure consumers (*10*). Some familiar questions raised by synthetic biology are not about safety, such as concerns about fair distribution of economic benefits and monopoly controls on inventions. The context is new, but we can build on earlier conceptual and practical work.

Other questions are thornier. Some scholars, journalists, and public interest groups are asking whether synthetic biology amounts to an overweening ambition to shape ourselves and other life forms (*1, 11*). We continue to see this question in the debates over assisted reproduction, genetic engineering, and surgical and pharmacological enhancement. It is at core a question about what it means to be human. It reemerges, it does not admit of crisp yes/no or good/bad answers, and we have not yet figured out how to best bring the concerns underlying this question into discussions of policy or the conduct of research—though we believe it is too soon to stop trying.

Bioethics does not need a new subfield to justify support for research on synthetic biology. Instead, we need to get better at appreciating and explaining that digging into familiar questions in new scientific contexts is the smartest way to inch forward.

The Hastings Center, Garrison, NY 10524, USA.

*Author for correspondence: parense@thehastingscenter.org

References and Notes

1. J. Boldt, O. Muller, *Nat. Biotechnol.* **26**, 387 (2008).
2. A. Balmer, P. Martin, *Synthetic Biology: Social and Ethical Challenges* (Biotechnology and Biological Sciences Research Council, Swindon, Wiltshire, UK, 2008); www.bbsrc.ac.uk/organisation/policies/reviews/scientific_areas/0806_synthetic_biology.pdf).
3. Synbiosafe E-conference, www.synbiosafe.eu/forum/.
4. E. Parens, J. Johnston, J. Moses, *Ethical Issues in Synthetic Biology: An Overview of the Debates* (Foresight and Governance Project at the Woodrow Wilson International Center for Scholars, Washington, DC, in press).
5. E. Parens, A. Asch, *Hastings Center Rep.* 29(4), S1 (1999).
6. M. S. Garfinkel, D. Endy, G. L. Epstein, R. M. Friedman, *Synthetic Genomics: Options for Governance* (J. Craig Venter Institute, with Center for Strategic and International Studies and Massachusetts Institute of Technology, Rockville, MD, 2007); www.csis.org/hs.
7. H. Bugl *et al.*, *Nat. Biotechnol.* **25**, 627 (2007).
8. S. M. Maurer, L. Zoloth, *Bull. At. Sci.* **63**(6), 16 (2007).
9. G. Church, *Nature* **438**, 423 (2005).
10. J. Macoubrie, *Informed Public Perceptions of Nanotechnology and Trust in Government* (Woodrow Wilson International Center for Scholars, Washington, DC, 2005).
11. ETC Group, *Extreme Genetic Engineering: An Introduction to Synthetic Biology* (ETC Group, Ottawa, Canada, 2007); www.etcgroup.org/upload/publication/pdf_file/602.

10.1126/science.1163821

DEVELOPMENTAL BIOLOGY

Return to the Proliferative Pool

Differentiated tracheal cells in the fruit fly larva paradoxically revert to a progenitor form as development advances to the adult stage.

Acaimo González-Reyes[1] and Jordi Casanova[2]

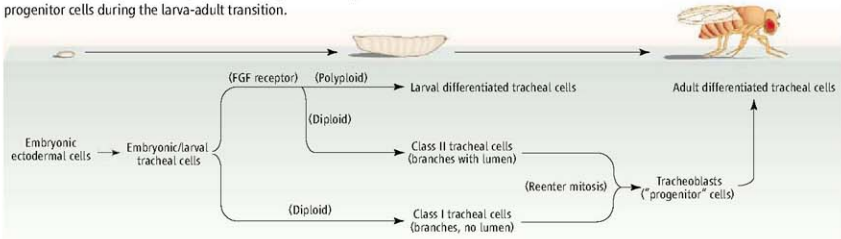
How do organs replace their differentiated and often highly specialized cells? Understanding the logic behind tissue homeostasis is essential for designing cell-regeneration–based therapies. On page 1496 in this issue, Weaver and Krasnow (*1*) report that certain structures of the adult tracheal system of the fruit fly *Drosophila melanogaster* originate from populations of larval progenitor cells, some of which were originally differentiated tracheal cells that regained the ability to proliferate.

form tubular branches, they lack an obvious lumen and do not form air transport tubes. Surprisingly, the other group of tracheoblasts (class II) originates from cells with a complex morphology that have undergone some differentiation, in that they have developed functional tubules in the larva.

Apart from these differences, both class I and class II cells share a number of features, such as reentry into cell division (mitosis) during the larval-adult transition and possession of small nuclei that, unlike other differentiated lar-

dedifferentiation process. If so, class II cells resemble the "facultative" stem cells described in vertebrate adult organs such as liver and pancreas, in which this particular type of stem cell, usually quiescent, can replenish and repair damaged tissue under certain physiological or pathological conditions (*4*, *5*). To better understand the biology of class II cells, a more detailed analysis of their behavior is needed. Does the acquisition of adult tracheal cell fates involve the asymmetric division of progenitor tracheoblasts, or do they divide

Tracheoblasts. The adult *Drosophila* tracheal system develops from committed cells in the larva that become progenitor cells during the larva-adult transition.



The *Drosophila* tracheal system is composed of tubular structures that transport oxygen to the different regions of the insect's body. Remarkably, the cellular mechanisms and genetic factors that underlie tracheal development are closely related to those operating in vertebrate vasculogenesis. During the generation of the adult tracheal system, almost all of the cells contributing to larval structures are eliminated and replaced by new ones. Where do these new adult tracheal cells come from? Weaver and Krasnow extend previous work (*2*, *3*) and show that some of them originate from the proliferation of two different, distant populations of larval tracheoblast (tracheal progenitors) (see the figure). One tracheoblast population (class I) arises from cells in which the embryonic-larval genetic program of tracheal development appears to arrest at an early step. Although these cells can

val cells, remain diploid and do not endoreplicate (multiplication of the genome without cell division). Whether these properties reflect a block in their tracheal differentiation remains to be determined, but other cells, which are morphologically and functionally related to class II cells, do endoreplicate and do not produce tracheoblasts. In addition, tracheoblasts derived from either class I or class II cells can give rise to different tracheal cell types. Although this may suggest that the larval tracheoblasts are multipotent progenitor cells, embryonic tracheal cells acquire their specific fate by virtue of their position within the tracheal branches and the intensity of a fibroblast growth factor (FGF) secreted from surrounding cells. Thus, the progeny of larval tracheoblasts may adopt diverse fates depending on their interaction with other tissues rather than on their history.

The ability of class II cells to generate tracheoblasts is particularly intriguing. These cells have advanced some steps into tracheal differentiation, and even though other interpretations are possible (*3*), their transformation into progenitors could be considered a

symmetrically to produce the cells that will make up adult structures? Is there a population of transit-amplifying cells (intermediate cells, between long-lived stem cells and their differentiated progeny, which can undergo limited rounds of division) that mediates the proliferation process?

In some vertebrate organs such as the liver and pancreas, facultative stem cells seem to be used only when other routes of tissue regeneration are not available (*6*, *7*). In the *Drosophila* adult tracheal system, the activity of tracheoblasts is a programmed response in development. Unless we consider tracheal metamorphosis as a form of stress response, the findings by Weaver and Krasnow reveal a new scenario in which cells initially enter differentiation with the certainty that they will return to a proliferative state later in development. This ability of differentiated cells to reenter proliferation may reflect an ancestral mode of metamorphosis used frequently in insects (*2*, *3*). It would be very interesting to determine whether the action of vertebrate facultative stem cells is limited to stress conditions and if they recapitu-

[1]Centro Andaluz de Biología del Desarrollo (CSIC-UPO), Sevilla, Spain. [2]Institut de Biologia Molecular de Barcelona (CSIC) and Institut de Recerca Biomèdica, Barcelona, Spain. E-mail: jcrbmc@cid.csic.es

late an earlier event in organ morphogenesis. Interestingly, the replacement in mammals of insulin-producing pancreatic cells in the adult requires preexisting cells (*7*), but it could also involve the reappearance of embryonic-like endocrine progenitor cells (*4*).

To become progenitor cells, class I and class II cells appear to escape the differentiation programs characteristic of more prosaic tracheal cells. This behavior could be determined by innate genetic instructions, by their cellular environment (much like the activity of a stem cell niche), or both. Mechanisms that maintain certain stem cell characteristics could be acting on specific cells to make them refractory to differentiation, as is the case in the murine embryonic stem cells (*8*). Regardless of the mechanism, "arrest" in the larval differentiation program to keep a cell as a potential progenitor cell appears to be a stepwise process. Thus, class II cells, capable of forming branches that transport air, can still behave as progenitor cells.

It is still unclear what establishes the "point of no return" after which a committed cell cannot revert into a progenitor cell. In the case of the *Drosophila* tracheal system, evidence points toward the triggering of endoreplication as a determining event (*1*–*3*), but many features of progenitor cell specification and activation in other systems remain to be elucidated. The study by Weaver and Krasnow is another excellent example of how the use of simpler, genetically tractable models such as the *Drosophila* tracheal system can aid in the interpretation of the genetic factors underlying progenitor cell biology in normal development or in stress conditions, an essential step for regenerative therapies.

References

1. M. Weaver, M. A. Krasnow, *Science* **321**, 1496 (2008).
2. M. Sato *et al.*, *Dev. Biol.* **318**, 247 (2008).
3. A. Guha, T. B. Kornberg, *Proc. Natl. Acad Sci. U.S.A.* **105**, 10832 (2008).
4. X. Xu *et al.*, *Cell* **132**, 197 (2008).
5. M. R. Alison *et al.*, *Cell Prolif.* **37**, 1 (2004).
6. E. L. Rawlins, B. L. M. Hogan, *Development* **133**, 2455 (2006).
7. Y. Dor, D. Melton, *Cell* **132**, 183 (2008).
8. Q.-L. Ying *et al.*, *Nature* **453**, 519 (2008).

10.1126/science.1163623

EVOLUTION

Dynamics of Body Size Evolution

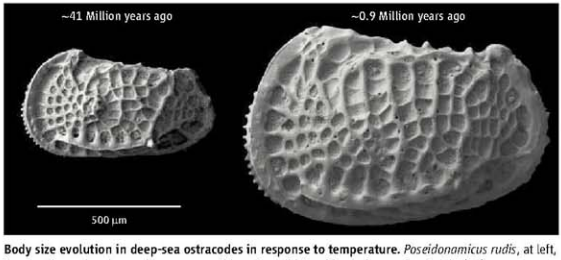
Is bigger better? Does climate affect size? The processes controlling body size evolution remain unclear.

Kaustuv Roy

Body size is one of the simplest organismic traits one can measure, yet it correlates with almost every aspect of the biology of a species, from physiology and life history to ecology. So, not surprisingly, biologists have long been interested in understanding how body size evolves. Two things are obvious when one looks at the distribution of body sizes of species within large groups: The sizes span multiple orders of magnitude, and species are not distributed uniformly within this range. Instead, most species tend to be small to intermediate in size, with few in the smallest and largest size classes. Thus, in most groups, size frequency distributions are skewed, even on a logarithmic scale, with the mode shifted toward smaller sizes. For example, living mammalian species range from about 2 g to 10^8 g with a modal size of about 100 g (*1*). Surprisingly, this bias toward smaller sizes persists despite a tendency for average size to increase over evolutionary time, a trend generally known as Cope's rule (*2*, *3*).

Models of body size evolution need to reconcile these two seemingly contradictory observations—a general tendency of size to increase over evolutionary time, yet the overall size frequency distribution staying biased toward small-bodied species. Two different types of evolutionary dynamics can lead to an increase in the average size of species over time. The first, Cope's rule in a strict sense, is a channeled increase in size where large species get larger and small ones go extinct (*3*). Alternatively, if groups arise near the small end of their size range—and paleontological data suggest that many do—then even random diffusion with a lower size limit increases the variance in size over time, leading to an increase in mean size (*2*). Reconciling such models with the shapes of empirical size frequency distributions is more difficult. Channeled increases in size obviously cannot produce a distribution that is biased toward smaller sizes. Similarly, even though stochastic models with a lower size bound can produce an increase in mean size over time, the resulting size distributions tend to be log-normal rather than the log-skewed distributions common in nature (*2*, *4*).

A recent model (*5*) provides one solution to this by making simple but elegant modifications to the multiplicative diffusion process. By incorporating a size-biased extinction rate and a strengthening of Cope's rule for the smallest species into a stochastic model, it successfully reproduced the size frequency distributions of mammal species. The model does make some key assumptions about size dependence of extinction and size change, but those seem well supported in mammals. More important, this model provides a general framework for modeling body size evolution that preserves insights from previous work (*2*, *4*) but also incorporates group-specific dynamics. It is too early to know whether the model is generally applicable; that would depend on whether it can predict size frequency distributions of groups such as marine mollusks, where neither extinction (*6*, *7*) nor Cope's rule (*3*) relate to size in the same way as in mammals.



Body size evolution in deep-sea ostracodes in response to temperature. *Poseidonamicus rudis*, at left, lived earlier and under much warmer conditions than did *Poseidonamicus major* at right (*12*).

PHOTO CREDIT: GENE HUNT/NATIONAL MUSEUM OF NATURAL HISTORY, WASHINGTON, DC

Section of Ecology, Behavior and Evolution, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093, USA. E-mail: kroy@ucsd.edu

Although phenomenological models are important for identifying key elements of body size evolution, they provide limited insights regarding the underlying processes. For example, if Cope's rule is indeed stronger for small mammals, then one has to ask why. Unfortunately, we are still far from such a process-based understanding of body size evolution, largely because of the complexity of the problem. Consider two generalizations about the connections among size, environment, and fitness that were suggested recently: "bigger is better" and "hotter is smaller" (*8*). The first is based on data from natural populations showing that larger individuals tend to have higher fitness. The second stems from observations that in laboratory-rearing experiments, higher temperatures generally result in smaller body sizes and also that species and individuals in cold climates are often larger than those in hotter areas, a trend known as Bergmann's rule.

Translating these "rules" into predictions about trajectories of size evolution is not straightforward. If bigger really is better, then we should have a world full of giants, yet most species are small. Clearly there are costs to getting bigger, which prevent a runaway Cope's rule. Such costs involve complex interactions among a multitude of factors including development time, population size, and patterns of resource use (*8*, *9*). In addition, the temperature-size rule suggests that the external environment, which changes in a complex and nonlinear manner over geologic time, is also important in driving size evolution. So, not surprisingly, simple process-based models of size evolution (such as one based on energetics) have not been widely accepted (*10*).

There is also the problem of scaling up from observations at the population level to macroevolutionary trends in size. The "bigger is better" rule is based on data from a few generations, and it is unclear whether it holds across geographically separated populations and macroevolutionary time. On the other hand, the temperature-size rule may indeed be relevant for macroevolution. Past climatic changes led to body size evolution consistent with the temperature-size rule in groups as disparate as woodrats (*11*) and deep-sea crustaceans (*12*) (see the figure). Furthermore, in some groups the temperature-size rule may have a relatively simple genetic basis; in the nematode *Caenorhabditis elegans*, it can be disrupted by a single nucleotide polymorphism (*13*).

Even though the processes governing body size evolution remain obscure, our collective actions are negatively affecting body sizes of many living species. Human exploitation of biological resources, from fisheries to forestry, is inherently size-selective where larger species and individuals are preferentially taken. As a result, body sizes of many species are much smaller now than, say, a century ago (*14*). Furthermore, abundances of large terrestrial and marine species are declining because of anthropogenic impacts, and many are threatened with extinction (*15*, *16*). Global warming may reinforce this trend toward smaller sizes through the temperature-size rule. In effect, then, our actions have set up a grand selection experiment where bigger is no longer better. Rapid microevolutionary responses to such selection have already been documented in laboratory experiments and in wild populations (*14*). Cope's rule is unlikely to be common in the future.

In a world where temperatures are rising and human exploitation of species is rampant, better understanding of ecological and evolutionary processes affecting body size is not simply an academic exercise; it is essential for effective management and conservation of species and ecosystems (*14*). The question now is not just why the world has so few giants, but how to keep the existing ones around for future generations.

References

1. T. M. Blackburn, K. J. Gaston, *Trends Ecol. Evol.* **9**, 471 (1994).
2. S. M. Stanley, *Evolution* **27**, 1 (1973).
3. D. Jablonski, *Nature* **385**, 250 (1997).
4. B. A. Maurer *et al.*, *Evolution* **46**, 939 (1992).
5. A. Clauset, D. H. Erwin, *Science* **321**, 399 (2008).
6. D. Jablonski, D. M. Raup, *Science* **268**, 389 (1995).
7. J. T. Smith, K. Roy, *Paleobiology* **32**, 408 (2006).
8. J. G. Kingsolver, R. B. Huey, *Evol. Ecol. Res.* **10**, 251 (2008).
9. J. H. Brown, B. A. Maurer, *Nature* **324**, 248 (1986).
10. C. R. Allen *et al.*, *Ecol. Lett.* **9**, 630 (2006).
11. F. A. Smith *et al.*, *Science* **270**, 2012 (1995).
12. G. Hunt, K. Roy, *Proc. Natl. Acad. Sci. U.S.A.* **103**, 1347 (2006).
13. J. E. Kammenga *et al.*, *PLoS Genet.* **3**, 358 (2007).
14. P. B. Fenberg, K. Roy, *Mol. Ecol.* **17**, 209 (2008).
15. K. J. Gaston, T. M. Blackburn, *Philos. Trans. R. Soc. London Ser. B* **347**, 205 (1995).
16. R. A. Myers, B. Worm, *Philos. Trans. R. Soc. London Ser. B* **360**, 13 (2005).

10.1126/science.1163097

CHEMISTRY

Bringing Stability to Highly Reduced Iron-Sulfur Clusters

A synthetic mimic of the most reduced iron-sulfur cluster in electron-transfer proteins shows a remarkable resemblance to protein-bound clusters.

Eckard Münck and Emile L. Bominaar

Many biochemical reactions are driven by electrons that are transferred to the reaction site from afar. Iron–sulfur clusters in proteins (*1*), including those with cuboidal Fe_4S_4 cores, can access different oxidation states and act as way stations for electrons; the oxidation state is designated by $[Fe_4S_4]^z$, where $z = 0, 1+, 2+, 3+$ is the formal core charge. In general, proteins use the (3+, 2+) or, most frequently, the (2+, 1+) redox couple. Evidence for the participation of the fully reduced $[Fe_4S_4]^0$ cluster in protein electron transfer has been scant, and a synthetic model in support of this oxidation state, as available for the higher oxidation states (*2*–*4*), has been lacking. Deng and Holm (*5*) have now provided such support in an innovative approach that replaces thiolates, used to simulate cysteinate binding in proteins, by electron-donating carbene ligands.

Some evidence supporting a role for the neutral (referred to as all-ferrous) cluster has come from one of the most intensely studied systems, namely nitrogenase from the bacterium *Azotobacter vinelandii*. Nitrogenase consists of two proteins: the molybdenum-iron protein (Av1), the locus of nitrogen reduction, and the Fe-protein (Av2), an electron transfer and effector protein. Av2 is a dimer of identical subunits that symmetrically coordinate a single Fe_4S_4 cluster through cysteine sulfurs (see the figure, left panel) (*6*). The Av2 dimer binds two molecules of MgATP (adenosine triphosphate), which are hydrolyzed in a coupled reaction that transfers electrons to Av1.

The accepted model for this electron transfer has been that Av2 uses the $[Fe_4S_4]^{2+,1+}$ redox couple. The electron transfer to Av1

Department of Chemistry, Carnegie Mellon University, 4400 Fifth Street, Pittsburgh, PA 15213, USA. E-mail: emunck@cmu.edu (E.M.); eb7g@andrew.cmu.edu (E.L.B.)

requires a conformational change in both proteins, made possible by hydrolysis of two ATP molecules per electron transferred (*7*). In 1995, it was discovered that the Av2 cluster could be reduced to the all-ferrous form (*8*). This discovery suggested that the Fe_4S_4 cluster of Av2 might function as a two-electron transfer agent, thereby decreasing the ATP requirement for nitrogen fixation by a factor of 2. However, subsequent studies of the 1+/0 redox couple suggested that its midpoint potential, −790 mV versus the normal hydrogen electrode, might be too low to be accessed in vivo (*9*).

Our understanding of iron-sulfur proteins has benefited from model clusters, which were synthesized by self-assembly using thiol ligands, in 1972 for $[Fe_4S_4]^{2+,1+}$ (*2*) and in 1985 for $[Fe_4S_4]^{3+}$ (*3*). However, thiolate-ligated $[Fe_4S_4]^0$ clusters are highly sensitive to oxidizing impurities and unsuitable for making the all-ferrous cluster. In 2005, Holm and co-workers reported a major advance with the isolation of an all-ferrous $[Fe_4S_4(CN)_4]^4$ cluster (*10*). Unfortunately, this complex could not be characterized in solution because it readily oxidized under these conditions.

Deng and Holm have now addressed the stability problem by replacing the cyanide ligand with an N-heterocyclic carbene ligand (1,3-diisopropyl-4,5-dimethylimidazole-2-xylidine, abbreviated as $Pr^i_2NHCMe_2$). In the all-ferrous cluster $[Fe_4S_4(Pr^i_2NHCMe_2)_4]$, neutral carbene ligands occupy the terminal positions at the tetrahedral iron sites. The x-ray structure of the complex (see the figure, right panel) shows that the carbene-ligated core has a cuboidal structure in which each bridging sulfide is bonded to three Fe^{2+} atoms.

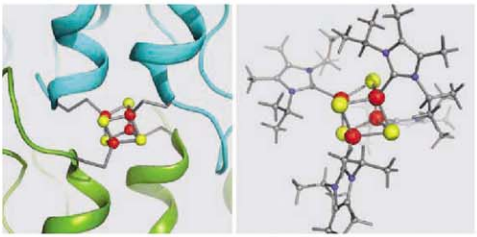
Although the carbene-ligated cluster is sensitive to oxidation, it can be manipulated in both the solid and solution states under ordinary anaerobic conditions. The utility of the carbene-substituted cluster as a model for the nitrogenase cluster depends on whether the properties of the $[Fe_4S_4]^0$ core are similar to cores in proteins stabilized by native thiolates. In this respect, data on the cluster's spin state, which is highly sensitive to its symmetry, are both encouraging and revealing.

The tetrahedral iron sites of the all-ferrous Fe_4S_4 cluster have four unpaired electrons yielding local spins $S = 2$ which, in turn, are coupled by exchange interactions via the bridging sulfides to give a ground state with cluster spin $S = 4$ (*11*). For a fairly symmetric model cluster with essentially equal exchange coupling constants, a simple analysis predicts a diamagnetic, $S = 0$, ground state. Thus, the observed $S = 4$ would represent a highly excited state, suggesting that the all-ferrous cluster must have lower symmetry.

Evidence for lower symmetry in the protein's all-ferrous cluster came from Mössbauer studies by Yoo *et al.* (*11*). Studies in applied magnetic fields revealed a pronounced 3:1 symmetry in the magnetic hyperfine parameters that was best explained by a spin structure with three iron spins aligned parallel to the cluster spin and one antiparallel, combining to total spin $S = 4$ for the cluster. In the absence of an applied field, the spectra revealed two quadrupole doublets with substantially different splittings in the same 3:1 ratio. This 3:1 core distortion was not recognized (*6*) from inspection of the 2.25 Å x-ray structure of all-ferrous Av2. The $S = 4$ state is observed when different reductants are used or when the $[Fe_4S_4]^{1+}$ cluster is radiolytically reduced at 77 K (*11*), which shows that the 3:1 distortion is not the result of cluster perturbations induced by the presence of reducing agents. Hans *et al.* recently reported a similar $S = 4$ state for the all-ferrous cluster of an activator protein from the bacterium *Acidaminococcus fermentans* (*12*).

Interestingly, Deng and Holm report a Mössbauer spectrum for the carbene cluster with a quadrupole pattern similar to that of Av2. Ongoing studies in our laboratory show that the carbene-substituted cluster also has an $S = 4$ spin state and that the magnetic hyperfine interactions in the synthetic and protein-bound clusters are very similar [see supporting online material (*13*)]. The observation of the same (distorted) electronic ground state in the cysteinate-bound protein cluster and the carbene-bound synthetic cluster suggests that the 3:1 pattern is an intrinsic and persistent property of the all-ferrous core, most likely rooted in the dependence of the exchange interactions on the structure of the core. In 2006, Lowery *et al.* reported (*14*) that the Av2 cluster can also attain a diamagnetic, $S = 0$ all-ferrous state at a surprisingly high potential (near −500 mV). This claim, however, has not yet been substantiated with a technique sensitive to the oxidation state of the cluster and, given the core's propensity toward symmetry breaking, should be viewed with caution.

There is still much to be learned about the all-ferrous Av2 cluster. The Deng-Holm model provides an x-ray structure with higher resolution than available for the protein-bound $[Fe_4S_4]^0$ cluster and thus offers a splendid opportunity to gain a deeper insight into the intrinsic relations between the molecular geometry and electronic structure of all-ferrous iron-sulfur clusters.



Natural and biomimetic core convergence. (Left) Close-up of the all-ferrous $[Fe_4S_4]^0$ cluster in the protein Av2, based on the 2.25 Å x-ray structure (*6*). The cluster is tethered to identical subunits (blue and green) by four cysteinate residues. **(Right)** All-ferrous cluster of $[Fe_4S_4(Pr^i_2NHCMe_2)_4]^0$ featuring carbenes as terminal ligands (*5*). The two cluster cores are essentially congruent. Color code: red (iron), yellow (sulfur), blue (nitrogen), and gray (carbon).

References and Notes

1. H. Beinert, R. H. Holm, E. Münck, *Science* **277**, 653 (1997).
2. T. Herskovitz *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **69**, 2437 (1972).
3. T. O'Sullivan, M. M. Millar, *J. Am. Chem. Soc.* **107**, 4096 (1985).
4. P. V. Rao, R. H. Holm, *Chem. Rev.* **104**, 527 (2004).
5. L. Deng, R. H. Holm, *J. Am. Chem. Soc.* **130**, 9878 (2008).
6. P. Strop *et al. Biochemistry* **40**, 651 (2001).
7. J. B. Howard, D. C. Rees, *Proc. Natl. Acad. Sci. U.S.A.* **103**, 17088 (2006).
8. H. C. Angove, S. J. Yoo, B. K. Burgess, E. Münck, *J. Am. Chem. Soc.* **119**, 8730 (1997).
9. M. Guo, F. Sulc, M. W. Ribbe, P. J. Farmer, B. K. Burgess, *J. Am. Chem. Soc.* **124**, 12100 (2002).
10. T. A. Scott, C. P. Berlinguette, R. H. Holm, H. C. Zhou, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 9741 (2005).
11. S. J. Yoo, H. C. Angove, B. K. Burgess, M. P. Hendrich, E. Münck, *J. Am. Chem. Soc.* **121**, 2534 (1999).
12. M. Hans, W. Buckel, E. Bill, *J. Biol. Inorg. Chem.* **13**, 563 (2008)
13. Supporting material is available on *Science* Online.
14. T. J. Lowery *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **103**, 17131 (2006).

Supporting Online Material

www.sciencemag.org/cgi/content/full/321/5895/1452/DC1
SOM Text
Fig. S1
Reference

10.1126/science.1163868

GEOLOGY

Understanding Soil Time

Efforts to maintain soils in a sustainable manner are complicated by interactions among soil components that respond to perturbation at vastly different rates.

Susan L. Brantley

Soils constitute the topmost layer of the regolith, the blanket of loose rock material that covers Earth's surface. An open system such as soil or regolith is sustainable, or in steady state, only when components such as rock particles are removed at the same rate they are replenished. However, soils are defined not only by rock particles but also by minerals, nutrients, organic matter, biota, and water. These entities—each characterized by lifetimes in regolith that vary from hundreds of millions of years to minutes—are often studied by scientists from different disciplines. If soils are to be maintained in a sustainable manner (*1*, *2*), scientists must develop models that cross these time scales to predict the effects of human impact.

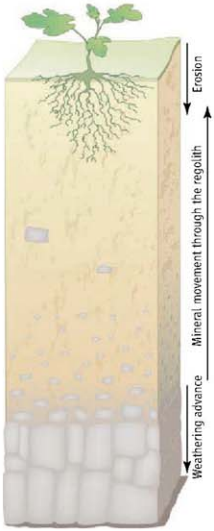
With respect to the longest time scales, geologists studying Earth's landscapes argue that continents have experienced balanced rates of tectonic uplift and erosion. Thus, the mass of rock particles produced by regolith-forming processes during uplift is balanced by the mass of particles eroded over geological time scales. At steady state, the mass of particles in the regolith "box," divided by the rate of removal of particles from that box, defines the particle residence time. If regolith is perturbed, the system moves toward a new steady state within a characteristic response time that—for linear systems—equals about 4 times the residence time.

These concepts are exemplified by observations of an undisturbed ridgetop in the Puerto Rican rainforest. At this site, the rate of particle mass loss due to dissolution and erosion (termed total denudation), cast as the rate of lowering of Earth's surface, is 0.04 mm/year (*3*). This rate is calculated by assuming that the rate of production of cosmogenic nuclides produced by penetration of cosmic rays into the upper 0.6 m of regolith is balanced by loss of these nuclides through denudation. The residence time for particles in this 0.6-m-thick "cosmogenic box" equals 15,000 years [= 0.6 m/(0.04 mm/year)]. If the thickness of this upper soil were perturbed, it would slowly return to its initial state over ~60,000 years.

As measured from cosmogenic isotopes, residence times in the upper 0.6 m of regolith range from 100 to 100,000 years for soils worldwide, depending on the intensity of tectonic activity (*4*). However, in the most tectonically quiescent areas of Africa, the cosmogenic technique no longer works, because the residence time of particles may reach hundreds of millions of years.

In contrast to geologists studying landscapes, geochemists interested in the chemical composition of the regolith focus on the response times of minerals. If one could stand on the 10-m-thick regolith at the ridgetop in Puerto Rico for a sufficiently long time, one would observe bedrock fragmenting into particles at 10 m depth that then diminish in size as they move upward and out of the regolith (see the figure). Quartz particles would ascend without disappearing, defining residence times similar to that of the rock particles. In contrast, feldspar dissolves from regolith particles during their trajectory across the lowest 30-cm layer of regolith, defining a residence time of 7500 years (*5*). More soluble minerals such as calcite can disappear even faster. Residence and response times of minerals, determined on the basis of chemical soil profiles, thus vary from hundreds of millions to hundreds of years.

If, instead of the minerals, the objects of study in the soil are the nutrients fixed from the atmosphere by organisms (*6*), the time scales of interest are generally shorter. Residence times of 100 to 1000 years are commonly estimated for soil organic matter, but some of this material turns over within 1 to 10 years (*7*, *8*). Residence times can be even shorter for nitrogen (*9*).



Inside the regolith. Weathering continuously replenishes the regolith while erosion removes soil at the surface. Minerals, organic matter, and water move through the regolith on different time scales, complicating efforts to define what sustainable soils are and how they could be maintained.

Scientists studying biota are often interested in the time scales that define how fast one ecosystem succeeds another after a disturbance. Generally, this response time is tens to hundreds of years. In fact, whether an ecosystem can ever reach steady state is a matter of debate. If it is possible, steady state is a complex function of the extent and frequency of disturbances such as fires and insect infestations (*10*).

The final component of soil considered here, water, responds at the shortest time scales. Water moves both downward (because of meteoric inputs) and upward (because of evapotranspiration mediated by roots that often extend to depths of tens of meters). Water residence times in regolith are measured with stable isotopes to decipher the interplay of "old" and "new" water. These water types are characterized by long or short residence times varying from tens of years to minutes.

When scientists within a discipline study soils, they generally focus on one of these time scales while ignoring faster and slower processes. Learning how soils will change in the future will require observations and models that cross time scales (*11*). For example, present-day and long-term denudation rates for catchments or soils have been shown to be equal across time scales in some cases, as required for sustainable soils. In other cases, the long-term and present-day denudation rates do not agree, perhaps because of variations in ecosystems, climate, glacial effects, extreme events, or human impact (*4*, *12*).

Another way to bridge time scales is to study chronosequences—soils formed on the same rock type in the same climate but for varying duration of weathering. For slow-weathering, undisturbed chronosequences,

Earth and Environmental Systems Institute, Pennsylvania State University, University Park, PA 16802, USA. E-mail: brantley@eesi.psu.edu

CREDIT: ADAPTED BY P. HUEY/SCIENCE

neither ecosystems nor regolith attain steady state; rather, they vary together as a result of the 30 or so bioessential elements mined by biota from rocks. Most important, phosphorus is extracted at depth by organisms, pumped upward, stored in biota and minerals, and recycled. Because phosphorus is lost to groundwater, however, depletion of regolith causes ecosystem degradation over 1000 to 10,000 years (*13*). Such coupled processes may be manifested in transformation of both above- and below-ground ecosystems as soils cross thresholds related to changes in pH, redox, and nutrient concentration (*14*, *15*). For example, subsurface ecosystems may become increasingly fungi-dominated as soils become phosphorus-limited (*13*).

The likelihood of crossing important thresholds is high today given the intensity of anthropogenic impact. Human activities have increased the long-term soil erosion rate by about a factor of 30 globally (*1*). Global agriculture has also caused nutrient depletion, especially in slow-weathering regions such as Africa. Largely to replenish nutrients, humans have doubled the input of fixed nitrogen into terrestrial ecosystems above prehuman values globally (*16*). The use of fertilizers replenishes soils but, given the time scale of soil water flow, also causes escape of nutrients and eutrophication in other ecosystems. For example, the transport of dissolved phosphorus from land to oceans has doubled, largely as a result of fertilizer use (*17*).

The need to maintain soils sustainably is now driving scientists to formulate models that describe not only how soil components react alone, but how they interact with each other in response to tectonic, climate, and anthropogenic forcing within the so-called Critical Zone—the zone extending from the depth of groundwater up to the outer limits of vegetation. Such models will provide the language that can allow scientists to communicate across disciplinary boundaries, but they must be tested across time scales with use of the sediment record, chronosequences, and observations of modern-day fluxes. Just as we use global climate models today to project future climate change, we will eventually be able to use global soil models to project future soil change.

References

1. B. H. Wilkinson, B. J. McElroy, *GSA Bull.* **119**, 140 (2006).
2. P. H. Bellamy *et al.*, *Nature* **437**, 245 (2005).
3. E. T. Brown, R. Stallard, M. C. Larsen, G. M. Raisbeck, F. Yiou, *Earth Planet. Sci. Lett.* **129**, 193 (1995).
4. F. Von Blanckenburg, *Earth Planet. Sci. Lett.* **242**, 224 (2006).
5. R. C. Fletcher, H. L. Buss, S. L. Brantley, *Earth Planet. Sci. Lett.* **244**, 444 (2006).
6. S. E. Trumbore, C. I. Czimczik, *Science* **321**, 1455 (2008).
7. J. S. Olson, *Ecology* **44**, 322 (1963).
8. K. Van Oost *et al.*, *Science* **318**, 626 (2007).
9. W. Parton *et al.*, *Science* **315**, 361 (2007).
10. M. G. Turner, W. H. Romme, R. H. Gardner, R. V. O'Neill, T. K. Kratz, *Landsc. Ecol.* **8**, 213 (1993).
11. A. F. White, S. L. Brantley, *Chem. Geol.* **202**, 479 (2003).
12. A. F. White, in *Kinetics of Water-Rock Interaction*, S. L. Brantley, J. D. Kubicki, A. F. White, Eds. (Springer, New York, 2008), pp. 463–494.
13. D. A. Wardle, L. R. Walker, R. D. Bardgett, *Science* **305**, 509 (2004); published online 17 June 2004 (10.1126/science.1098778).
14. J. A. Wiens, *Funct. Ecol.* **3**, 385 (1989).
15. O. A. Chadwick, J. Chorover, *Geoderma* **100**, 321 (2001).
16. P. M. Vitousek, H. A. Mooney, J. Lubchenco, J. M. Melillo, J. M. Melillo, *Science* **277**, 494 (1997).
17. G. M. Filippelli, in *Phosphates: Geochemical, Geobiological, and Materials Importance*, M. J. Kohn, J. Rakovan, J. M. Hughes, Eds. (Mineralogical Society of America, Washington, DC, 2002), pp. 391–425.

10.1126/science.1161132

GEOLOGY

An Uncertain Future for Soil Carbon

A detailed knowledge of how carbon cycles through soils is crucial for predicting future atmospheric carbon dioxide concentrations.

Susan E. Trumbore and Claudia I. Czimczik

Predictions of how rapidly the large amounts of carbon stored as soil organic matter will respond to warming are highly uncertain (*1*). Organic matter plays a key role in determining the physical and chemical properties of soils and is a major reservoir for plant nutrients. Understanding how fast organic matter in soils can be built up and lost is thus critical not just for its net effect on the atmospheric CO_2 concentration but for sustaining other soil functions, such as soil fertility, on which societies and ecosystems rely. Recent analytic advances are rapidly improving our understanding of the complex and interacting factors that control the age and form of organic matter in soils, but the processes that destabilize organic matter in response to disturbances (such as warming or land use change) are poorly understood.

There is broad agreement on the major pathways of the soil carbon cycle (see the figure). Plants are the main source of carbon to soils through tissue residues or via root exudates and symbiotic fungi. These inputs are broken down, transformed, and respired by soil fauna and microorganisms. Some of the carbon converted into microbial biomass and by-products is in turn converted into new microbial biomass ("recycled") (*2*). Some organic molecules, such as pyrogenic compounds, may accumulate because of recalcitrance. However, most soil organic matter consists of relatively simple molecules that organize through interactions with surfaces and with each other (*3*). Organic matter persists in soil mainly because it is physically isolated from decomposition by microbes—for example, by incorporation into aggregates (*4*) or sorption into mineral (or other organic) surfaces (*5*, *6*). On balance, nearly all the carbon that enters soil as plant residues each year either decomposes and returns to the atmosphere or is leached from soils within a few decades to centuries.

The rates of accumulation and loss of soil carbon are estimated from two kinds of information: direct observations of changes in the amount of organic matter, and inferences based on the age of organic matter as measured by radiocarbon. These rates vary dramatically depending on the time scale of observation, and they reflect differences in the dominant processes contributing to the stabilization of organic matter.

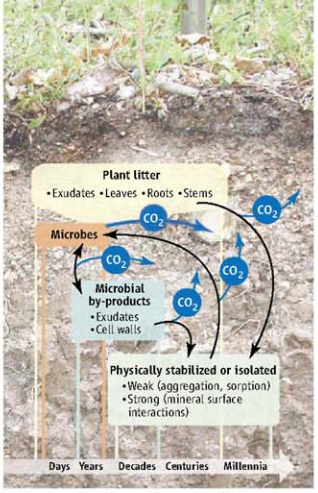
On time scales of months to years, observed rates of mass loss during decomposition of fresh plant litter nearly balance rates of plant litter addition to soils (~2 to 10 Mg C ha^{-1} year^{-1}). Litter decomposition is thus the major pathway for loss of carbon from soils (see the figure), and rates are controlled by factors such as litter quality, soil faunal and microbial community composition, and climate (*7*).

On millennial time scales, changes in carbon stocks cannot be observed directly. They are estimated by comparing carbon storage at carefully selected sites that differ in the time since bedrock weathering started (soil age) but are similar in other soil-forming factors such as bedrock material, climate, and vegetation. Such comparisons yield rates of change in soil

Department of Earth System Science, University of California, Irvine, CA 92697, USA. E-mail: setrumbo@uci.edu; czimczik@uci.edu

carbon stores of ~0.02 Mg C ha^{-1} year^{-1}, much slower (by a factor of 100 to 500) than fresh litter decomposition (*8*); on these time scales, the amount and age of soil carbon are controlled by changes in mineral surfaces related to weathering (*9*).

Most of the concern associated with soil carbon response to global change involves organic carbon stocks that can change over decades to centuries. Changes in these kinds of organic matter are too small to be observed over a few years, and on millennial time scales they are obscured by other factors—such as vegetation productivity and nutrient supply (*10*)—that vary with soil mineralogy. Our understanding of carbon dynamics on these intermediate time scales relies either on quantifying changes in carbon stocks and stable carbon isotopes after disturbances such as fire or land use change, or on following the incorporation of radiocarbon produced in the 1960s by nuclear weapons tests into soil carbon pools.



Carbon transformation pathways in soil. The scale at the bottom indicates the mean age of organic carbon typically found in each pool. Blue arrows indicate CO_2 production during transformation from one pool to another.

Such measurements identify substantial stores of soil carbon that can accumulate and be lost at intermediate rates (~0.1 to 10 Mg C ha^{-1} year^{-1}). Processes that can stabilize or destabilize organic carbon on these time scales include alterations to the quantity, age, and quality of plant litter inputs; shifts in the makeup, spatial distribution, and function of soil fauna and microbial communities; alteration of weak stabilization processes such as aggregate formation; and changes in mineral surfaces under altered redox or pH conditions. Such processes will respond on decadal time scales to changes not only in climate, but also in nutrient deposition or vegetation. To predict future concentrations of atmospheric CO_2, it is critical to better understand how much carbon is vulnerable to destabilization on decadal to centennial time scales, and which processes provide the most important controls for a given ecosystem.

The shortcomings of the current understanding are apparent when trying to predict the response of mineral soil carbon stores to global warming. A good example is the debate over the temperature dependence of decomposition rates for different carbon pools (*11*). In general, fresh plant material decomposes faster at higher temperatures (*7*). However, it is less clear whether or how carbon stabilized on mineral surfaces responds to temperature changes. Conceptual models like that depicted in the figure are too simplistic in treating production and decomposition as separate rather than linked processes (*12*). Changes in temperature will influence all parts of the soil-plant system; simple temperature functions may provide a means to average across this complexity, but are not likely to be useful for predicting responses outside observed conditions. Progress will require temperature manipulations of whole ecosystems, coupled with observations of soil carbon and isotope fluxes in concert with modeling (*13*).

The most robust predictions of future soil carbon change involve accelerated decomposition of relatively fresh plant material that persists because of flooding or freezing conditions, rather than by interactions with minerals. For example, high-northern-latitude regions that store vast amounts of carbon in relatively undecomposed forms, and where temperatures are rising faster than the global mean, are predicted to become net carbon sources to the atmosphere over the next century, because decomposition rates increase more than plant productivity does (*14*).

Outside of these special cases, a number of outstanding issues still limit our ability to predict soil carbon response. For example, plant residues arrive in the soil with different ages (years for leaves, centuries for tree stems). Hence, the radiocarbon age of soil organic matter is not merely a measure of the time period organic matter spends in soils, which may bias interpretations of its stability. Soils are not well-mixed media, and the timing of degradation and stabilization processes is also regulated by the complex spatial distribution of organic matter, microorganisms, and minerals (*5*). Sampling that integrates over that spatial domain may mix very young and old components to arrive at an average that is not adequate for describing rates of response to short-term change.

Most detailed studies of soil carbon age and chemistry are conducted on small plots for a few years, yet processes operating at larger spatial scales over decades to centuries (such as erosion, fire, nutrient deposition, or vegetation change) may ultimately determine the impact of soils on atmospheric CO_2. For example, fire-dominated Mediterranean and boreal ecosystems accumulate surface litter between burning events. Increasing burned area in a given year can return carbon faster to the atmosphere than it accumulates in unburned areas, making the region a net carbon source (*15*). Rapidly changing land-use patterns, as observed in the tropics, can be more important for evaluating soil carbon balance than are the factors causing variable rates of carbon loss or gain in an individual field (*16*). Such landscape-scale processes are crucial for the global carbon budget but are only beginning to be addressed in field studies or ecosystem carbon models.

Future progress will come from studies that combine measures of microbial community and activity, soil physics and chemistry, and the structure, age, and chemical nature of organic matter stored in and exiting soils. These studies should not focus only on improving models of the upper 10 to 20 cm of mineral soil at one location, but must recognize that soil processes extend in three dimensions, as deep as roots and across landscapes. Continued changes in climate will ultimately show how soil carbon will respond, but predicting changes would be the safer route given the importance of soil organic matter in sustaining society.

References

1. P. Friedlingstein *et al.*, *J. Clim.* **19**, 3337 (2006).
2. G. Gleixner *et al.*, *Org. Geochem.* **33**, 357 (2002).
3. I. Kögel-Knabner, *Soil Biol. Biochem.* **34**, 139 (2002).
4. J. Six *et al.*, *Plant Soil* **241**, 155 (2002).
5. K. Ekschmitt *et al.*, *Geoderma* **128**, 167 (2005).
6. M. Kleber *et al.*, *Biogeochemistry* **85**, 9 (2007).
7. W. Parton *et al.*, *Science* **315**, 361 (2007).
8. W. H. Schlesinger, *Nature* **348**, 232 (1990).
9. M. S. Torn *et al.*, *Nature* **389**, 170 (1997).
10. S. L. Brantley, *Science* **321**, 1454 (2008).
11. E. A. Davidson, I. A. Janssens, *Nature* **440**, 165 (2006).
12. M. Heimann, M. Reichstein, *Nature* **451**, 289 (2008).
13. Y. Luo, *Annu. Rev. Ecol. Evol. Syst.* **38**, 683 (2007).
14. K. Dutta *et al.*, *Global Change Biol.* **12**, 2336 (2006).
15. J. W. Harden *et al.*, *Global Change Biol.* **6**, 174 (2000).
16. K. W. Holmes *et al.*, *Biogeochemistry* **74**, 173 (2005).

10.1126/science.1160232

Cooling, Heating, Generating Power, and Recovering Waste Heat with Thermoelectric Systems

Lon E. Bell

Thermoelectric materials are solid-state energy converters whose combination of thermal, electrical, and semiconducting properties allows them to be used to convert waste heat into electricity or electrical power directly into cooling and heating. These materials can be competitive with fluid-based systems, such as two-phase air-conditioning compressors or heat pumps, or used in smaller-scale applications such as in automobile seats, night-vision systems, and electrical-enclosure cooling. More widespread use of thermoelectrics requires not only improving the intrinsic energy-conversion efficiency of the materials but also implementing recent advancements in system architecture. These principles are illustrated with several proven and potential applications of thermoelectrics.

Broad societal needs have focused attention on technologies that can reduce ozone depletion, greenhouse gas emissions, and fossil fuel usage. Thermoelectric (TE) devices, which are semiconductor systems that can directly convert electricity into thermal energy for cooling or heating or recover waste heat and convert it into electrical power, are increasingly being seen as having the potential to make important contributions to reducing CO_2 and greenhouse gas emissions and providing cleaner forms of energy.

In this Review, I compare solid-state TE systems with more familiar mechanical providers of heating, cooling, and electrical power generation, such as air conditioners, refrigerators, heat pumps, heat exchangers, and turbine engines. Both classes of devices are similar in that they employ a working fluid—in TEs, this is electrical current, whereas in conventional closed-system heat engines steam or freon substitutes are the common working fluids. These classes have complementary regimes in which they can provide good performance.

Solid-state energy conversion has great appeal in terms of its simplicity as compared with systems that must compress and expand a two-phase (gas/liquid) working fluid. However, except in a limited number of cases, the operational efficiencies of TE systems have fallen short of the targets needed for them to be used broadly. Nevertheless, several commercial uses have been realized, including thermal cycles for DNA synthesizers, car seat cooler/heaters, laser diode coolers, and certain low-wattage power generators. Successful applications have capitalized on the small size of these devices, their robustness in demanding environments, or their rapid response time.

Two important pathways will lead to additional applications for TEs. One will be improving the intrinsic efficiencies of TE materials, and many efforts are underway to accomplish this. Another is to improve the way in which existing TEs are currently used, which will be the main focus of this overview.

BSST, Irwindale, CA 91706, USA.

Thermoelectrics As Heat Engines

TE devices are solid-state heat engines. Unlike today's air conditioners, which use two-phase fluids such as the standard refrigerant R-134A, TE devices use electrons as their working fluid. Figure 1 demonstrates the principal effects that govern their performance.

In 1834, Peltier observed that if a current is applied across a junction of dissimilar electrically conductive materials, either heating or cooling can occur at the junction. When the current is reversed, the opposite effect is observed. Figure 1A illustrates why this occurs. Electric current is propagated by electrons in n-type materials and by holes (traveling in the opposite direction) in p-type materials, be they semiconductors, metals, or semimetals. If voltage is applied in the right direction across a p-n junction, electron/hole pairs are created in the vicinity of the junction. Electrons will flow away from the junction in the n-type material, and holes will flow away in the p-type material. The energy to form them comes from the junction region, cooling it. On the opposite end, electrons and holes stream toward junctions where pairs recombine. This process releases energy and heats the junctions. At the bottom of Fig. 1 is a typical TE module, configured so that all junctions on one side heat and those on the other side cool.

In 1821, Seebeck noticed that the needle of a magnet is deflected in the presence of dissimilar metals that are connected (electrically in series and thermally in parallel) and exposed to a temperature gradient. The effect he observed is the basis for TE power generation. As shown in Fig. 1B, if the junctions at the top are heated and those at the bottom are cooled (producing a temperature differential), electron/hole pairs will be created at the hot end and absorb heat in the process. The pairs recombine and reject heat at the cold ends. A voltage potential, the Seebeck voltage, which drives the hole/electron flow, is created by the temperature difference between the hot and cold ends of the TE elements. The net voltage appears across the bottom of the TE element legs. The Seebeck effect forms the basis of the operation of TE couples (thermocouples) used extensively in temperature-measurement systems. Electrical connections can be made from the TEs to an external load to extract power.

In order for this process to be efficient, it is necessary to find materials that are good electric conductors, otherwise electron scattering generates heat on both sides of the barrier and throughout the materials. Also, the materials must be poor thermal conductors, otherwise the temperature difference that must be maintained between the hot and cold sides will produce large heat backflow. Similarly, the Seebeck effect should be maximized. Optimization of these three parameters is compromised because all three are affected by the electronic properties of the materials. Because the working fluid (electrons) conducts unwanted heat as well as electric current, and the Seebeck effect decreases as the electrical conductivity increases, it is necessary to optimize these properties simultaneously (*1*). The highest performance is achieved with heavily doped semiconductors, such as bismuth telluride or silicon germanium. Finally, for semiconductors, it is desirable to have a base material that can be both p- and n-type–doped, so that the same material system can be used on both sides of the junctions.

It is useful to compare the electrical current as a working fluid with the gas/liquid two-phase fluids in conventional air conditioners. The key difference that allows a refrigeration system in a building to achieve up to 60% of the maximum theoretical efficiency (as compared with 12% for TEs to date) is that cooling and heat-rejection components can be physically well separated, and large temperature differences do not lead to the high heat backflow that penalizes efficiency in TE systems.

Practical Thermoelectric Devices

In a working TE device, segments of p-type– and n-type–doped semiconductor materials, such as suitably doped bismuth telluride, are connected by shunts to form an electric circuit. The shunts are made of an excellent electrical conductor, such as copper. A voltage drives a current through the circuit, passing from one segment to another through the connecting shunts. For determining efficiency, this configuration is equivalent to the electrons passing directly from one TE material to the other. Conventional TE cooling/heating modules are constructed of pairs of TE segments, repeated about 100 times, and organized into arrays like the one shown in Fig. 1. When current

flows within the module, one side is cooled and the other heated. If the current is reversed, the hot and cold sides reverse also. The geometry for power generators (Fig. 1B) is conceptually the same. In this case, the top side is connected to a heat source and the bottom to a heat sink. TE power generators often are similar in physical form to cooling modules except that fewer taller and thicker elements are used.

A figure of merit, ZT, expresses the efficiency of the p-type and n-type materials that make up a TE couple. The parameter Z is the square of the Seebeck voltage per unit of temperature, multiplied by the electrical conductivity and divided by the thermal conductivity, and T is the absolute temperature. In today's best commercial TE cooling/heating modules, ZT is about 1.0, and in air-conditioning applications is about one-quarter as efficient as a typical conventional system, such as one that uses R-134A. Ideal TE system efficiency increases nonlinearly with ZT, so that to double efficiency, ZT has to increase to about 2.2. To achieve a fourfold increase (to equal the efficiency exhibited by today's two-phase refrigerants), ZT would need to increase more substantially to about 9.2.

To maximize power-generation efficiency, ZT should be as high as possible, and the temperature differential between the hot and cold sides should be as large as possible. The material properties that make up Z vary with temperature, so that materials exhibit optimum performance over a relatively narrow temperature range. As a result, in order to maximize the efficiency of power-generation modules, individual TE elements are usually formed from two and sometimes three different TE materials laminated together in the direction of current flow to form segmented elements. Each TE material in the laminate structure is chosen to have superior performance over the range of its temperature exposure (*2*). For effective waste heat recovery from vehicle exhaust (an operating condition with about a 350°C temperature differential), the efficiency needs to be about 10% and the corresponding average ZT should be about 1.25 to increase mileage up to 10%. For primary power generation, the net efficiency needs to be about 20% and have an average ZT of 1.5 or greater (at an 800°C temperature differential).

Given that the merits of solid-state energy conversion are easily understood and accepted, why are TE devices not more broadly used? The principal reason is that efficiency has been too low to be economically competitive. Use in large-scale air conditioners for homes and commercial buildings falls into this category. Besides low efficiency, a second reason is that the cost of traditional TE modules per watt of cooling, heating, or power generation has been too high to allow the displacement of existing technologies, with the exception of those few applications in which the beneficial characteristic of being solid state outweighs cost and performance limitations. A lack of design knowledge and design tools has created yet another barrier to broader usage. In part, this is to be expected because the technology has a legacy of low efficiency and high cost, and the small number of applications has limited the development of commercially designed software.

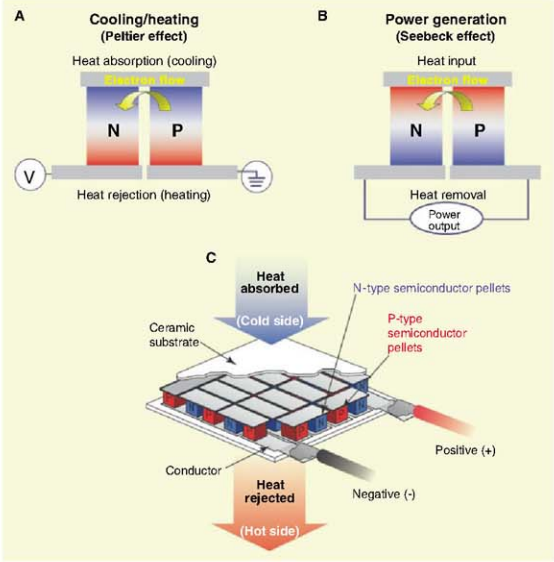


Fig. 1. TE heat engines. **(A)** When current is run across a TE junction, it heats or cools through the Peltier effect, depending on the direction of the current flow. **(B)** When heat flows across the junction, electrical current is generated through the Seebeck effect. **(C)** Practical TE generators connect large numbers of junctions in series to increase operating voltage and spread heat flow.

Improving Performance

In 1993, the U.S. government's Office of Naval Research and Defense Advanced Research Projects Agency asked interested researchers to propose pathways to improve ZT for cooling and heating applications (*3*). A specific interest was to determine whether the then-emerging nanotechnology and its potential quantum-scale synthesis could lead to new superior TE materials. In 1993, Hicks and Dresselhouse published a theoretical model predicting the effect on ZT of confining electrons to two-dimensional quantum wells (*4*). They calculated that the Seebeck coefficient could be increased and the thermal conductivity could be suppressed. The promise of this concept and other ideas from within the TE community led the U.S. government to fund several innovative approaches in the mid-1990s. This initiative set in motion a substantial increase in both theoretical and TE-material developmental research. By 2001, Venkatasubramanian of Research Triangle Institute announced achievement of a room-temperature ZT of about 2.4 for a nanoscale structure made by alternating layers of two TE materials that both enhanced the Seebeck coefficient and suppressed thermal conductivity (*5*). The next year, Harman of Lincoln Laboratory published results claiming a ZT of up to 3.2 at about 300°C for a material with nanoscale inclusions that dramatically reduced thermal conductivity (*6*). In 2003, Kanatzidis at Michigan State University led a team in the development of a complex bulk tertiary material with a ZT of at least 1.4 at 500°C (*7*). Recently, Heremans at Ohio State University and an international team claimed reaching a ZT of 1.5 at 500°C (*8*). Despite these promising results, efficiency gains at the device level have yet to be demonstrated. The scaling of the nanomaterials has proven to be quite difficult and is still in the development stage. The bulk material has yet to be made commercially available.

The theoretical and nanoscale experimental results led to increased interest in pursuing applications enabled by the promised material advancements. At the same time as the material gains were forthcoming, attempts were made to identify other opportunities to achieve system-level performance gains. In total, four sources for these opportunities

can be identified. The first is an obvious one: Increase the material *ZT*. The second is to look to continuous improvement in design optimization and to reduce parasitic losses by using nontraditional materials in device fabrication. These opportunities lead to efficiency gains of up to 25%; the potential for large gains is limited because most early uses of TE devices were for aerospace applications and great efforts had already been devoted to design optimization.

The system is suited for use with a working fluid, such as air and water. Flow patterns are similar to those of counterflow heat exchangers. However, in this case, the heat transported from one fluid to the other is modified by the TE engines as it passes through the system. In cooling/heating mode, the TE elements boost the heat quality so that one of the opposing fluid streams is heated and the other is cooled. For the conditions illustrated, the efficiency can be about double that of a single module operating with all elements at the same temperature (*10*). Figure 2B shows a similar geometry operating in power-generation mode. In this case, system efficiency gains are about 30% greater than those of TE heat engines in which the incoming working fluid is combusted without being preheated by the waste side of the TE array (*11*). The cycles can be combined with higher-*ZT* materials to compound the performance gains.

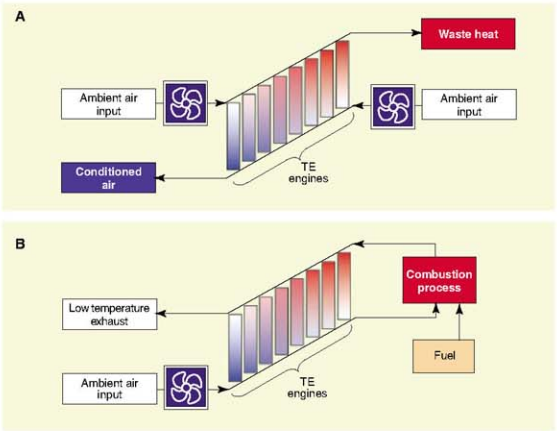


Fig. 2. Thermodynamic cycles. By optimizing each element along the thermal gradient, the engine resembles a gas turbine engine (the high-efficiency Brayton cycle) rather than the less efficient diesel cycle, in which the temperature and pressure conditions of every element (TE junction or combustion cylinder) are the same. This approach is shown for heating and cooling in (**A**) and for power generation in (**B**).

The third is to determine whether alternative thermodynamic cycles could be used to improve efficiency. In thermodynamic terms, each p and n TE couple (Fig. 1) is a separate heat engine and, in principle, could operate independently of the other engines that make up a TE device. If each engine could operate optimally (that is, at the ideal temperature and current), system-level efficiency could increase. The analog is to compare the efficiency of two common heat engines that burn oil. In a diesel engine, each cylinder is an independent heat engine, but all cylinders operate at the same temperature and pressure conditions, the diesel cycle. The engine delivers about 30 to 45% efficiency. In contrast, in a turbine engine, such as is used in municipal electric power–generation systems, every stage of the compressor and expansion sections operates optimally for the working-fluid conditions at each point. This is a regenerative Brayton cycle, and efficiency in modern systems is about 60 to 65%, nearly double that of the diesel cycle. We developed a cycle analogous to the Brayton cycle, in which the TE engines are arranged as shown in Fig. 2A (*9*).

The fourth opportunity comes from the realization that there is no length scale in the equations that determines the efficiency of a TE engine. Thus, the amount of TE material used in the construction of a perfect, theoretically lossless TE engine is arbitrary. In real devices, system performance does degrade as the device is made smaller, because the relative impact of parasitic electrical and thermal-loss mechanisms increases as size decreases. Also, manufacturing tolerances and electrical isolation requirements place practical limitations on device size. An alternative stack TE configuration in Fig. 3B has lower parasitic losses from the electrical connections (shunts) between TE elements than does the traditional configuration shown in Fig. 3A. Maximum benefit occurs if the stack design is used in combination with a reduction in the electrical resistance at the TE-material/shunt interfaces.

Other parasitic losses tend to increase when less TE material is used, so TE material volume can be lowered to the point where the benefits of the reduced resistance inherent in the stack geometry equal the increased losses from other sources. Under many practical operating conditions, the weight of TE material used can be reduced by a factor of 6 to a factor of 25 (*12*). By employing the design technology behind this configuration, TE material costs for high-capacity systems are generally much lower. For example, TE cooling and heating systems with the traditional configuration are cost-competitive up to about 400 thermal watts of output, but increases to about 4000 W with the stack design. Power generators have been restricted to uses in harsh, remote environments where reliability justifies higher costs. With substantial reductions in TE material usage, a broader spectrum of commercial applications becomes economically viable.

Present Applications

TE-powered devices have been in production since the bismuth telluride–based room-temperature materials were developed in the late 1950s. As with many innovative technologies that offer new functionality (in this case, refrigerators and heat engines the size of matchboxes), military applications developed first. TEs are used to produce –80°C temperatures to operate the sensors in infrared imaging systems for heat-seeking missiles and night-vision systems. The development of silicon germanium high-temperature power-generation materials led directly to the production of heat engines for space applications with no moving parts that could operate in the absence of sunlight. Solar cells, another type of heat engine, are effective and can be used as far as the orbit of Mars, but beyond that distance the solar radiant flux is not adequate to power spacecraft. All power sources for U.S. and former-USSR deep-space probes have used TE heat engines to convert heat generated by nuclear fissile material to electricity (*13*).

More-recent applications take advantage of lowered costs and greater yields made possible by adopting semiconductor manufacturing processes for fabricating TE materials and devices. TE-powered devices are now in mass production for cooling, heating, and temperature-control applications in several important markets. Miniature TE modules keep laser diodes at constant temperature to stabilize operating wavelengths (*14*). Biological assaying has been revolutionized by the development of polymerase chain reaction (PCR) systems, and PCR systems use TE devices to thermally cycle microliter quantities of enzymatic reactions through exact series of temperature cycles (*15*). The process is used to multiply specific sequences of DNA material for analytical testing purposes.

Climate-control seat (CCS) systems have been developed to provide rapid seat cooling in the summer and equally fast heating in the winter (*16*). The CCS is being installed in more than 500,000 vehicles a year. It increases passenger comfort and is beginning to be used to provide a degree of comfort when the vehicle engine is off. Fuel consumption is reduced in hybrid vehicles in hot driving conditions because the engine can be turned off when the vehicle is coasting, braking, or stopped.

Portable beverage and picnic coolers were an early commercial application that combined the small size, light weight, and electric operation that characterize TEs to open a new market for ice-free portable cooling systems (*17*). Product offerings have since expanded to include coolers that are quiet and vibration-free. They have proven popular for replacing traditional vapor-compressor refrigeration in wine-storage cabinets, hotel room mini-refrigerators, and office water coolers.

circuits allow and hence create a hazard, which now can be averted while giving the user added heating power and the benefits of cooling.

Microprocessor electronic TE-based cooling systems have seen limited application as add-ons to boost personal-computer clock speeds. Present TE systems do not demonstrate sufficient performance gains for acceptance by the general personal-computer market. However, TE cooling of small electronic enclosures, such as those used to cool the various low-power computer boards that control industrial equipment, is efficient and cost-effective. The systems have a long history of successful (if limited) application where reliability is critical and cooling capacity can be limited to less than 1000 W (*19*).

Present TE power-generation systems have been limited to uses for which their durability and maintenance-free operation dominate other performance criteria. Examples of important uses are power for remote data communication systems for

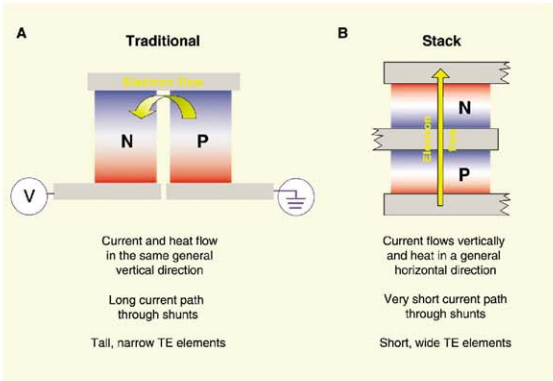


Fig. 3. Alternative TE junction geometries. **(A)** A traditional junction. Current and heat flow in the same general direction, and there is a long current path through shunts and tall narrow TE elements. **(B)** A stack junction. The current flow is perpendicular to the heat flow, the current path is minimized through shunts, and the TE elements are short and wide.

Personal temperature-control systems that provide cooling as well as heating for the office environment have also come onto the market. A desktop unit uses the high-efficiency thermodynamic cycle to condition a person's immediate environment while keeping input electrical power to that of a personal computer and within industry limits for power consumption by office furnishings and electrical outlets (*18*). For this and many other applications, the TE systems, when used in the heating mode, operate as heat pumps and not simply as resistive heaters. Efficiency is about two and a half times that of the traditional office resistance heaters. Usually, traditional heaters draw more electrical power than the ratings for cubicle oil and gas pipelines, polar weather station power generators, and cathodic protection for oil drilling platforms (*20*). TE generators are chosen for these applications because of their proven reliability (often maintenance-free operation for 20 years), durability under extreme conditions, and very little if any performance degradation over their operating life.

Future Applications

The separate technological advances in materials, cycles, and power density can be combined readily to compound benefits. More-efficient cycles are coming into use in automotive, electronic enclosure, and personal climate-control applications. As higher-*ZT* materials become available, they should be able to be incorporated into existing TE designs with relatively little modification. Performance will be enhanced for all of these uses to the extent that conventional devices can be replaced with no operating-cost increase. However, the most exciting prospects for TE technology are that new uses will be enabled that have beneficial impacts on the environment.

If the average *ZT* reaches 2, room, home, and commercial solid-state heating, ventilating, and air-cooling systems become practical. The systems would replace R-134A, which has a greenhouse gas equivalence of 1430 times that of CO_2 (*21*), with electric current as the working fluid. Not only would the systems have zero CO_2 equivalency from leakage and refrigerant disposal, but they would have exceptional heat-pumping performance, so that if the source of electrical power were green, fossil fuel usage would be eliminated in the winter as well.

Recently published review articles give a comprehensive overview of trends and accomplishments in TE material developments (*22*–*25*). Other recent publications claim gains in bulk material *ZT* of up to 40% (*26*) and advancements in nanostructured silicon (a poor TE material in bulk form) to performance levels at the nanoscale on par with those of today's best commercial materials (*27*, *28*). Several of the new TE materials are grown as 5- to 30-μm-thick films (*5*, *6*, *29*, *30*). Standard semiconductor fabrication methods are used to form these materials into arrays of submillimeter couples. The resulting devices exhibit thermal response times as short as 20 μs and can cool, heat, or generate electric power. In addition to integrated circuit hot-spot cooling, these properties suggest possible new applications requiring temperature change for function, such as fast DNA analysis on a nanoliter scale, continuous environmental or hazard assaying, real-time monitoring of complex biological processes, and control and power supplies for remote-sensing systems.

Because of their ruggedness, portability, and ready ability to be electrically powered, TE systems should provide more-efficient and better-performing temperature control in vehicles of many types, including cars, trucks, trains, and aircraft. The advantages, in addition to eliminating unfriendly refrigerants, would again be that the very efficient cooling and heating would be contained in the same package and operate with the same controls. At a *ZT* of 2, cooling and temperature control of microprocessors, communications circuitry, electro-optical systems, and other electronic components become attractive. The clock speed and operating life of many chip circuits decreases rapidly with increased temperature, so that effective thermal management becomes beneficial on several counts (*31*).

Average *ZT* in the range from 1.5 to 2 would enable substantial waste-heat harvesting and primary power-generation applications. Various government-sponsored programs are underway in the United States and Japan to increase vehicle

mileage by converting a fraction of the heat in the exhaust systems of trucks and cars to electric power (*32*, *33*). The power would be available to drive power steering, brakes, water pumps, turbochargers, and other vehicle subsystems electrically. Less electrical power would need to be produced by the alternator, which has efficiency losses caused by the load these subsystems put on engines today. On average, the electrical subsystems weigh less and can be positioned more favorably in the vehicle away from the engine, so that secondary but still important system-level gains occur. All of these factors combine to improve mileage and reduce costs. Estimates vary depending on the degree of system integration and on driving conditions, but the U.S. Department of Energy target of 10% fuel reduction appears to be within reach at these higher ZT levels (*34*).

Gains of about 5 to 10% would be possible in diesel-powered cogenerators that are becoming widely used for onsite power generation in developed countries and for 5000- to 20,000-W primary generators in developing countries. In another proposed cogenerator concept, the solar spectrum is split into shorter wavelengths that yield high photovoltaic-conversion efficiency and longer wavelengths that heat a TE generator (*35*). Recent studies by Pacific Northwest National Laboratory suggest that industrial waste-heat recovery in aluminum smelting, glass manufacture, and cement production is practical at a ZT of 2 (*36*). At the same ZT, it appears possible to replace small internal combustion engines such as those used in lawn mowers, blowers, and small outboard motorboats with external combustion TE engines.These engines would be very quiet and nearly vibration-free. They could burn a wide spectrum of fuels, such as propane, butane, liquified natural gas, and alcohols, and would not necessarily depend on fossil oil as a fuel source.

Outlook

Until recently, TE technology has languished despite the astonishing gains made in electronics, photonics, and other solid-state fields. Now, 15 years after U.S. government initiatives spurred resurgence in TE research, substantial progress is evident. More-efficient thermodynamic cycles and designs that reduce material costs are coming into commercial production. If the final enabling advancement, higher ZT in TE materials, is realized, gas-emission–free solid-state home, industrial, and automotive air conditioning and heating would become practical. In power generation, fuel consumption and CO_2 emissions would be reduced by electric power production from vehicle exhaust. Industrial waste-heat recovery systems could reduce emissions by providing supplemental electrical power without burning additional fossil fuel. The question is, Is TE technology on a path to overcome the historic limitations of low efficiency and high cost per watt of power conversion that have limited its applications in the past? If so, TE solid-state heat engines could well play a crucial role in addressing some of the sustainability issues we face today.

References and Notes

1. S. W. Angrist, *Direct Energy Conversion* (Allyn and Bacon, Boston, MA, 1965), chap. 4, pp. 144–150.
2. G. L. Snyder, *Appl. Phys. Lett.* **84**, 2436 (2004).
3. C. B. Vining, paper presented at the European Conference on Thermoelectrics, Odessa, Ukraine, 10 to 12 September 2007.
4. L. D. Hicks, M. S. Dresselhause, *Phys. Rev. B* **47**, 16631 (1993).
5. R. Venkatasubramanian *et al.*, *Nature* **413**, 597 (2001).
6. T. C. Harman *et al.*, *Science* **297**, 2229 (2002).
7. K. F. Hsu *et al.*, *Science* **303**, 818 (2004).
8. J. P. Heremans *et al.*, *Science* **321**, 554 (2008).
9. L. E. Bell, paper presented at the 21st International Conference on Thermoelectrics, Long Beach, CA, 25 to 29 August 2002.
10. R. W. Diller, Y. Chang, paper presented at the 21st International Conference on Thermoelectrics, Long Beach, CA, 25 to 29 August 2002.
11. L. E. Bell, paper presented at the 22nd International Conference on Thermoelectrics, Hérault, France, 17 to 21 August 2003.
12. L. E. Bell, paper presented at the 23rd International Conference on Thermoelectrics, Adelaide, Australia, 25 to 29 July 2004.
13. R. D. Abelson, *Thermoelectrics Handbook* (CRC Press, Boca Raton, FL, 2006), chap. 56, pp. 1–29.
14. Marlow Industries, www.marlow.com.
15. Global Medical Instrumentation; product details are at www.gmi-inc.com/BioTechLab/mjresearchptc100.htm.
16. J. Lofy, L. E. Bell, paper presented at the 21st International Conference on Thermoelectrics, Long Beach, CA, 25 to 29 August 2002.
17. Igloo; product details are at www.igloo-store.com/product_detail.asp?T1=IGL+COOL+CHILL+40+BL&HDR=t&.
18. "Blowing Hot and Cold," *Forbes*, www.forbes.com/leadership/2008/01/15/temperature-office-workplace-lead-careers-cx_tw_0115bizbasics_slide_3.html?thisSpeed=15000.
19. G. S. Mikalauskis, "Selecting a Thermoelectric Cooler," www.electronicproducts.com/ShowPage.asp?Filename=teca.may2004.html.
20. Global Thermoelectric, "Thermoelectric Generators for Cathodic Protection," www.farwestcorrosion.com/fwst/dcpower/global01.htm.
21. P. Forester *et. al.*, in *Climate Change 2007: The Physical Science Basis, Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change* (Cambridge Univ. Press, New York, 2007), p. 212.
22. F. DiSalvo, *Science* **285**, 703 (1999).
23. B. Sales, *Science* **295**, 1248 (2002).
24. G. J. Snyder, E. Toberer, *Nat. Mater.* **7**, 105 (2008).
25. T. Tritt, M. A. Subramanian, *MRS Bull.* **31**, 188 (2006).
26. B. Poudel *et al.*, *Science* **320**, 634 (2008).
27. A. Hochbaum *et al.*, *Nature* **451**, 163 (2008).
28. A. Boukai *et al.*, *Nature* **451**, 168 (2008).
29. H. Boettner *et al.*, *Thermoelectrics Handbook* (CRC Press, Boca Raton, FL, 2006), chap. 46, pp 1–18.
30. J. Bowers *et al.*, *Material Research Society Fall Meeting* (Material Research Society, Boston, MA, 2007) **1044**, p. U10-06.
31. A. Jain, S. Ramanathan, paper presented at the Intersociety Conference on Thermal and Thermomechanical Phenomena in Electronic Systems, San Diego, CA, 30 May to 2 June 2006.
32. J. Bass, N. B. Elsner, A. Leavitt, paper presented at the 13th International Conference on Thermoelectrics, Kansas City, Missouri, 30 August to 1 September 1994.
33. K. Ikoma, M. Munekiyo, K. Furuya, T. Izumi, K. Shinohara, paper presented at the 17th International Conference on Thermoelectrics, Nagoya, Japan, 24 to 28 May 1998.
34. L. E. Bell, paper presented at the 21st International Conference on Thermoelectrics, Long Beach, CA, 25 to 29 August 2002.
35. T. Tritt, H. Bottner, L. Chen, *MRS Bull.* **33**, 366 (2008).
36. Industrial Technologies Program, "Engineering Scoping Study of Thermoelectric Generator Systems for Industrial Waste Heat Recovery" (U.S. Department of Energy, 2006).

37. The author is president of BSST and owns stock in the company, and BSST owns patents related to this work.

10.1126/science.1158899

A Mutation in Hairless Dogs Implicates *FOXI3* in Ectodermal Development

Cord Drögemüller,[1] Elinor K. Karlsson,[2] Marjo K. Hytönen,[3,4] Michele Perloski,[2] Gaudenz Dolf,[1] Kirsi Sainio,[3] Hannes Lohi,[4] Kerstin Lindblad-Toh,[2,5] Tosso Leeb[1]*

Mexican hairless dogs (xoloitzcuintles) were considered sacred by the Aztecs, and statues of the dogs date back 3700 years. Hairless dogs were described as *Canis aegyptius* by Linné (*1*). The phenotype of hairless dogs is now classified as canine ectodermal dysplasia (CED) because these dogs have missing or abnormally shaped teeth in addition to a hair coat that is sparse or absent (Fig. 1). CED is inherited as a monogenic autosomal semidominant trait (*2*) linked to chromosome 17 (CFA 17) (*3*) in Mexican and Peruvian hairless dogs and in Chinese crested dogs. Heterozygous dogs are hairless, and homozygous mutants die during embryogenesis.

The spontaneous CED mutation offers an opportunity to gain further insights into the development of ectodermal appendages such as hair and teeth. Mutations in the *EDA*, *EDAR*, and *EDARADD* genes of the ectodysplasin signaling pathway produce a closely related phenotype called anhidrotic ectodermal dysplasia. Anhidrotic ectodermal dysplasia is characterized by the absence of hair, dentition abnormalities, and alterations in certain exocrine glands, the latter feature seen not at all or less severe in CED (*4*).

We used genomewide association mapping (*5*) to identify the causative genetic mutation (*6*). One single-nucleotide polymorphism (SNP) on CFA 17 showed strong association with the hairless phenotype. Haplotype association testing revealed an 8-SNP, 160-kb haplotype with perfect phenotypic concordance. We fine-mapped the associated haplotype in a larger set of Chinese crested dogs along with Mexican and Peruvian hairless dogs. The mutant chromosomes of all hairless dogs from the three breeds shared identical marker alleles in a 102-kb interval between positions 41,045,331 and 41,147,100 on CFA 17 (*6*).

In this region of the dog genome (CanFam2.0), only two genes were predicted, including a dog-specific processed pseudogene with a truncated open reading frame. The other predicted gene, LOC483075, showed homology to the forkhead box (FOX) transcription factor gene family. In mammalian genomes, this family consists of about 50 members encoding several important regulators of development.

The dog genome assembly contains a gap in the region of the first exon of LOC483075, and we determined the complete sequence of this region (*6*). The resultant previously uncharacterized gene showed highest similarity to the mouse *Foxi3* gene (LOC232077), and we propose the symbol *FOXI3* for the dog gene (*7*). A human *FOXI3* gene with conserved open reading frame can be found in Celera's human genome assembly, whereas a putative sequencing error is present in the extremely GC-rich parts of the human reference assembly (build 36.3). To examine whether *FOXI3* might be a reasonable functional candidate gene for CED, we analyzed its expression in mouse embryos. Whole-mount in situ hybridization experiments revealed highly specific expression of *Foxi3* transcripts in the epithelium of the hair and whisker placodes and developing teeth, consistent with a functional role for the gene in the development of these organs (*6*).

Sequence analysis of the *FOXI3* gene in hairless and coated Chinese crested dogs revealed a 7-bp duplication within exon 1 (Fig. 1) that produces a frameshift and a premature stop codon. We found perfect concordance of the duplication with the CED phenotype in 140 hairless dogs from the three studied breeds. The duplication was absent from 55 coated animals from the studied breeds and 32 coated dogs from 19 other breeds (*6*). Although it is unclear whether the mutant protein is actually expressed, with more than 95% of the normal FOXI3 protein missing, it is unlikely that the mutant protein is functional. Any mutant protein produced will contain 218 altered amino acids and could potentially interfere with normal cellular function in a dominant-negative mechanism. However, we think the most likely explanation for the CED phenotype in hairless dogs is haploinsuffiency of FOXI3 caused by the 7-bp duplication. According to this hypothesis, a complete lack of FOXI3 in homozygous mutants causes the embryonic lethality.

We speculate that FOXI3 is a downstream target of the ectodysplasin signaling pathway. This study illustrates how the extreme morphological diversity of dogs can be harnessed to gain new insights into developmental biology.

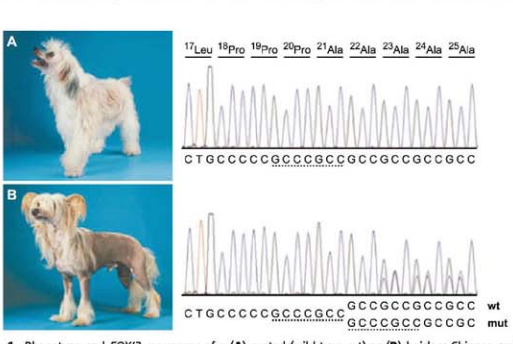


Fig. 1. Phenotype and *FOXI3* sequences of a **(A)** coated (wild-type, wt) or **(B)** hairless Chinese crested (mut) dog. Hairless dogs have a 7-bp duplication (underlined) in exon 1 of *FOXI3* (c.57_63dup7).

References and Notes

1. C. von Linné, *Systema Naturae*... (Halae Magdeburgicae, ed. X, 1758), p. 39.
2. R. Robinson, *J. Hered.* **76**, 217 (1985).
3. D. P. O'Brien *et al.*, *J. Hered.* **96**, 727 (2005).
4. T. Kimura, *Exp. Anim.* **45**, 1 (1996).
5. E. K. Karlsson *et al.*, *Nat. Genet.* **39**, 1321 (2007).
6. Materials and methods are available as supporting online material on *Science* Online.
7. T. Ohyama, A. K. Groves, *Dev. Dyn.* **231**, 640 (2004).
8. Supported by the UniBern Research Foundation, the Academy of Finland, and the Sigrid Juselius Foundation.

Supporting Online Material

www.sciencemag.org/cgi/content/full/321/5895/1462/DC1
Materials and Methods
SOM Text
Figs. S1 to S5
References and Notes

30 June 2008; accepted 14 August 2008
10.1126/science.1162525

[1]University of Berne, 3001 Berne, Switzerland. [2]Broad Institute of Harvard and Massachusetts Institute of Technology, Cambridge, MA 02142, USA. [3]Institute of Biomedicine, 00014 Helsinki, Finland. [4]University of Helsinki, 00014 Helsinki, Finland. [5]Uppsala University, 75124 Uppsala, Sweden.

*To whom correspondence should be addressed. E-mail: Tosso.Leeb@itz.unibe.ch

Enhanced Sensitivity of Photodetection via Quantum Illumination

Seth Lloyd

The use of quantum-mechanically entangled light to illuminate objects can provide substantial enhancements over unentangled light for detecting and imaging those objects in the presence of high levels of noise and loss. Each signal sent out is entangled with an ancilla, which is retained. Detection takes place via an entangling measurement on the returning signal together with the ancilla. This paper shows that for photodetection, quantum illumination with m bits of entanglement can in principle increase the effective signal-to-noise ratio by a factor of 2^m, an exponential improvement over unentangled illumination. The enhancement persists even when noise and loss are so great that no entanglement survives at the detector.

The conventional way to detect the presence of an object is to shine light in its direction and to see whether any is reflected back. If the object is far away, only a small percentage of the light will be reflected to a detector. If the object is immersed in noise and thermal radiation, then whatever light is reflected must be distinguished from the noisy background. In the case of quantum bits, it is known that the sensitivity of estimation processes can be enhanced by entangling a signal qubit with an ancilla and by making an entangling measurement on the returning qubit together with that ancilla (*1*–*3*). Entanglement and squeezing are known to enhance various aspects of amplification and imaging (*4*–*6*). This paper shows that entanglement can in principle give an enhancement of sensitivity for photon counting. For example, the photons can be entangled via frequency, as in the output from a spontaneous parametric downconverter. The intuition is that, if the signal is entangled with an ancilla, then it is harder for noise to masquerade as the returning signal. This intuition turns out to be correct, even though noise and loss completely destroy the entanglement between signal and ancilla at the detector. In fact, the entanglement-induced enhancement holds only in the presence of noise and loss. This resistance to noise and loss stands in stark contrast with existing methods that use entanglement to enhance the precision of measurement, where noise and loss rapidly destroy any enhancement [for a review of such methods, see (*7*)]. Moreover, the entanglement-induced enhancement of sensitivity is substantial: It grows exponentially with the number of bits of entanglement between signal and ancilla.

This paper presents the simplest possible mathematical treatment of quantum illumination, where signal and ancilla consist of individual photons. The entangled states used correspond to the single-photon sector of the output of a parametric downconverter, and the measurements required can be performed by upconversion. Two models are analyzed, one a low-noise model and the other with high levels of noise: Both exhibit exponential enhancement of sensitivity. Multiphoton signals and ancillae that exhibit coherence between states with different photon number are likely to surpass the bounds derived here.

Suppose that a single photon is sent at the object to be detected. If the object is not there, the signal received at the detector consists only of thermal and background noise. Even if the object is there, the majority of time the photon is lost, and again only noise is received. Once in a while, however, the photon is reflected back in a perturbed form. The dynamics corresponding to this situation can be modeled as beam splitter with reflectivity, η, that mixes in the vacuum with the signal state, followed by thermalization, which injects noise with average photon number, b, into each optical mode (e.g., frequency or polarization). No object corresponds to a beam splitter with $\eta = 0$. The presence of an object corresponds to a beam splitter with a very small nonzero reflectivity.

To make the analysis tractable, the following simplifying assumptions are made. In the first model analyzed here, the number of noise photons per mode is taken to be small: $b \ll 1$. This scenario corresponds, for example, to single photons in the optical regime directed at a distant target that is bathed in thermal radiation at temperature significantly below optical energies. The detector can distinguish between d modes per detection event: $d = WT$, where W is the bandwidth of the detector and T is the temporal length of the detection window. The time window for detection is sufficiently small that at most one noise photon is detected at a time, i.e., $db \ll 1$. Additional, nonthermal noise can be tolerated as well, as long as fewer than one photon arrives per detection event. These assumptions, small reflectivity and small numbers of noise photons, are made for ease of analysis only. Quantum illumination can also work in the presence of high numbers of noise photons and high detection rates. In the second noise model analyzed below, arbitrarily high amounts of noise can be tolerated.

First, consider the case of unentangled light. Send a single photon in the state ρ toward the region where the object might be. The two different dynamics corresponding to object there and object not there are as follows (*8*):

Case (0), object not there: $\rho \rightarrow \rho_b \otimes \ldots \otimes \rho_b$, where ρ_b is the thermal state of a mode with b photons on average, and $\otimes$ is the tensor product. Because the average number of photons bd received per detection event is much less than 1, the thermal state can be approximated as

$$\rho_0 = (1 - db)|vac\rangle\langle vac| + b\sum_{k=1}^{d} |k\rangle\langle k| \quad (1a)$$

Here, $|vac\rangle$ is the vacuum state of the modes, and $|k\rangle$ is the state where there is a single photon in mode k and no photons in the other modes. Because $\| \rho_0 - \rho_b \otimes \ldots \rho_b \|_1 = (db)^2 + O[(db)^3]$, the exact thermal state can be safely replaced by ρ_0 as long as we are evaluating the expression to lowest order in db.

Case (1), object there: $\rho \rightarrow (1-\eta)\rho_0 + \eta\rho_{th}$ where ρ_{th} is the thermalized version of ρ. It is straightforward to verify that $\|\rho - \rho_{th}\|_1 = db + O[(db)^2]$. Accordingly, if we are not interested in terms of $O(\eta db)$, then we can safely approximate $(1-\eta)\rho_0 + \eta\rho_{th}$ as

$$\rho \rightarrow \rho_1 = (1 - \eta)\rho_0 + \eta\rho \quad (1b)$$

Now repeatedly send single photons in the state $|\psi\rangle$ to detect presence or absence of the object. The single-shot minimum probability of error is obtained by projecting onto the positive part of $\rho_1 - \rho_0$ [see (*1*–*3*) for a treatment of discrimination between different dynamics in the case of qubits]. This measurement consists simply of verifying whether a returning photon is in the state $|\psi\rangle$ or not. If the measurement yields a positive result, then we guess that the object is there. If the result is negative, then we guess that the object is not there. The conditional probabilities of the outcomes "yes" and "no" given the presence or absence of the object are

$$\begin{aligned} p(\text{no}|\text{not there}) &= 1 - b \\ p(\text{no}|\text{there}) &= (1 - b)(1 - \eta) \\ p(\text{yes}|\text{not there}) &= b \\ p(\text{yes}|\text{there}) &= b(1 - \eta) + \eta \end{aligned} \quad (2)$$

If we iterate the optimal single-shot measurements, Eq. 2 determines the number of trials required to reveal presence or absence of the object. The number of trials required depends on the ratio η/b. If $\eta/b > 1$, then a received photon is more likely to be a signal photon than a noise photon: the signal-to-noise ratio is greater than 1. We'll call this the good regime. In the good regime, the number of trials required to detect the object, if there, goes as $O(1/\eta)$: one simply sends photons

W. M. Keck Center for Extreme Quantum Information Processing (xQIT), Department of Mechanical Engineering, MIT 3-160, Cambridge, MA 02139, USA. E-mail: slloyd@mit.edu

until one receives one back. If the object is there, then one receives a photon back considerably before one would expect a photon given thermal photons alone.

Similarly, if $\eta/b < 1$, then a received photon is more likely to be a noise photon than a signal photon: The signal-to-noise ratio is less than 1. We'll call this the bad regime. In the bad regime, most of the photons received are noise photons. Here, one must count photons until one can separate the thermal distribution with b photons on average from the distribution when the object is there, which has $b + \eta$ photons on average. By using the usual formulae for sampling the outcomes of Bernoulli trials, one finds that it takes $O(8b/\eta^2)$ photons on average to distinguish between presence or absence of the object in the bad regime.

These estimates of the number of trials required to distinguish between the presence and the absence of the object are just lower bounds because they come from iterating the optimum single-shot measurement. The true asymptotic limit on the number of trials is given by the quantum Chernoff bound (*9*, *10*). The quantum Chernoff bound for photocounting is calculated in (*11*) and confirms that the iteration of the single-shot measurement is indeed optimal.

The error probabilities for detecting presence or absence of the object do not depend on the number of signal and detector modes, d. The number of modes doesn't matter because all modes other than the one in which the photon is sent are in a thermal state. These other modes give us no information about the presence or absence of the object. To detect the object, we need only monitor the mode in which we sent the photon to see whether more than the expected number of photons come back.

Now look at the effect of entanglement on our ability to detect the object. Construct the entangled state $|\psi\rangle_{SA} = (1/\sqrt{d})\sum_k |k\rangle_S |k\rangle_A$ for signal photon and ancilla photon and send the signal photon to toward where the object is likely to be. This state can be created, for example, by taking the output of a spontaneous parametric downconverter in the low-photon number regime, matching its time-bandwidth product to the time-bandwidth product of the detector, and selecting out the one signal photon/one idler photon sector [photodetection at the receiver automatically postselects this state (*11*)]. The signal photon is sent off, and the idler photon is retained as the ancilla.

The two different dynamics corresponding to object or no object now take a different form because the state of the ancilla must be included. If the signal photon is lost, the ancilla photon goes to the completely mixed state:

Case (0), object not there:

$$|\psi\rangle_{SA}\langle\psi| \rightarrow \rho_{SA0} = \rho_0 \otimes \frac{I_A}{d} =$$

$$[(1-db)|vac\rangle_S\langle vac| + bI_S] \otimes \frac{I_A}{d} + O(b^2) \quad (3a)$$

Case (1), object there:

$$|\psi\rangle_{SA}\langle\psi| \rightarrow \rho_{SA1} = (1-\eta)\rho_{SA0} + \eta|\psi\rangle_{SA}\langle\psi| + O(b^2, \eta b) \quad (3b)$$

Here I_S and I_A are the identity operators on the single photon Hilbert spaces for the system and ancilla, respectively. As before, the single-shot minimum error probability is obtained by projecting onto the positive part of $(\rho_{SA1} - \rho_{SA0})$, which in turn simply corresponds to determining whether any returning photon is in the state $|\psi\rangle_{SA}$.

Making the optimal single-shot measurement and evaluating the conditional error probabilities for the entangled case yields

$$p_e(\text{no}|\text{not there}) = 1 - \frac{b}{d}$$

$$p_e(\text{no}|\text{there}) = (1 - \frac{b}{d})(1 - \eta)$$

$$p_e(\text{yes}|\text{not there}) = \frac{b}{d}$$

$$p_e(\text{yes}|\text{there}) = (1 - \eta)\frac{b}{d} + \eta \quad (4)$$

Comparison with the unentangled case, Eq. 2, immediately reveals that the effect of entanglement is to reduce the effective noise from b to b/d. This reduction reflects the fact that in the entangled case a noise photon together with the fully mixed ancilla is d times less likely to be confused for a signal photon entangled with the ancilla than a noise photon is likely to be confused with a signal photon in the unentangled case. Entanglement reduces the effective signal to noise by a factor of d. Again, the single-shot minimum error probability for the entangled case can be shown to coincide with the asymptotic minimum error probability by evaluating the quantum Chernoff bound (*11*).

Comparing the error probabilities for entangled states, as given by Eq. 4, with those for unentangled states, Eq. 2, we see that there are once more two regimes. The good regime now occurs when $\eta d/b > 1$. In the good regime, it once again takes $O(1/\eta)$ trials to determine whether the object is there. Comparing the entangled case to the unentangled case above, we see that the number of trials is the same in the good regime in both cases, but the good regime extends d times further in the entangled case than in the unentangled case, where the good regime occurred for $d/b > 1$.

The extension of the good regime via the use of entangled photons can be understood as follows. As before, the quantum Chernoff bound analysis given in (11) shows that the optimal detection strategy is to measure any incoming photon together with the ancilla to see whether the two photons are in the state $|\psi\rangle_{SA}$. If the photon that returns is the signal photon, then it will pass the test. If the photon that returns is a noise photon, then the ancilla is in the state I_A/d, and the noise photon together with fully mixed ancilla are d times less likely to be found in the state $|\psi\rangle_{SA}$. A noise photon in the entangled case is d times less likely to pass the test and be confused as a signal photon than a noise photon in the unentangled case. In other words, the presence of entanglement makes it d times harder for a noise photon to masquerade as a signal photon. Entanglement effectively enhances the signal-to-noise ratio by a factor of d.

The bad regime for the entangled case occurs for $\eta d/b < 1$. In this case, the number of iterations required to determine whether the object is there or not is $O(8b/\eta^2 d)$. Comparing with the unentangled bad regime, we see that the entangled bound is d times better than the unentangled bound: Quantum illumination reduces the number of trials needed to detect the object by a factor d. Entanglement effectively enhances the signal-to-noise ratio by the degree of entanglement, even in the bad regime.

The fact that entanglement yields an enhancement in the bad regime is particularly interesting because in the bad regime the combination of noise and loss ensures that no entanglement between signal and ancilla survives, an effect that also appears in the qubit case (*1*–*3*). Nonetheless, even though signal and ancilla are unentangled at the detector, a noise photon still finds it d times harder to masquerade as a signal photon entangled with an ancilla photon. Entanglement effectively enhances the signal-to-noise ratio of detection by a factor of d, where d is the number of entangled modes. Measured in terms of m, the number of e-bits of entanglement, the enhancement is $d = 2^m$. The enhancement is exponential in the number of e-bits.

The noise model above assumed small numbers of noise photons per mode and a sufficiently short detection window that at most one photon is detected during that window. These assumptions can be relaxed. Consider, for example, a noise model in which if the object is not there, then the state received at the detector is $\rho_S = I_S/D$, where D is the total number of possible noise states that possess the same energy as the signal. In other words, the noise is described by the microcanonical ensemble, and $D = 2^m$, where m is the entropy of the noise measured in bits. If the object is there, then with a small probability η the signal is reflected back unchanged, whereas with probability $1 - \eta$ the signal is replaced by noise. This model corresponds, for example, to the detection of a reflective, multifaceted object tumbling through space. Most of the time, one receives only background noise, but every once in a while a facet reflects the signal back to the detector.

This "microcanonical" noise model can be analyzed by using the same tools used above, including the quantum Chernoff bound. As shown in (*11*), the model yields similar conclusions to the low-temperature model analyzed above. When an unentangled state $|\psi\rangle$ is sent, then there are two regimes for detection. When $\eta D > 1$, we are in the good regime, with signal-to-noise ratio greater than one, and $O(1/\eta)$ signals must be sent to detect the presence of the object. When $\eta D < 1$, we are in

the bad regime, with signal to noise less than one, and $O(8/D\eta^2)$ signals must be sent. By contrast, when half of a fully entangled state is sent out, the good regime is extended to the case $\eta D^2 > 1$: Just as in the low-noise case above, entanglement extends the good regime by a factor D. In the bad regime, where $\eta D^2 < 1$, the number of signals that must be sent to detect the object is $O(8/D^2\eta^2)$: Just as in the low-noise case, initial entanglement reduces the number of signals that needs to be sent by a factor of $D = 2^m$, where m is the number of bits of entanglement. As before, in the bad regime, no entanglement survives at the receiver.

This example elucidates nature of the enhancement afforded by entanglement. For the microcanonical noise model, the sensitivity of the unentangled detection goes as the dimension of the signal space, whereas the sensitivity of the entangled detection goes as the dimension of the signal space squared. This enhancement is similar to that obtained in superdense coding (*12*), in which entanglement allows one to use a single qudit (signal space with dimension D) to send two classical dits (dimension D^2). In the low-noise model, by contrast, the linear nature of the noise model rendered the sensitivity of detection independent of the number of signal modes. For both noise models, adding entanglement enhances the sensitivity by the dimension of the signal space.

Quantum illumination is a potentially powerful technique for performing detection, in which signal is entangled with an ancilla and entangling measurements are made at the detector. Entanglement enhances the effective signal-to-noise ratio because a noise photon has a harder time masquerading as an entangled signal photon compared with a noise photon masquerading as an unentangled signal photon. The enhancement of sensitivity and effective signal-to-noise ratio that quantum illumination provides is exponential in the number of bits of initial entanglement and persists even in the presence of large amounts of noise and loss, when no entanglement survives at the receiver. This entanglement-induced enhancement for detection is reminiscent of entanglement-assisted communication capacity (*12*, *13*), where large enhancements in the presence of noise also occur.

Quantum illumination can be used for ranging and imaging: The effective signal-to-noise enhancement persists as we sweep the detection window in time to determine the distance to the object and as we sweep the signal beam in space to image features of the object. Many practical questions remain; notably, can the requisite entangled measurements be performed efficiently? Detection of entangled photons is a well-studied topic [see, e.g. (*14*)], and the requisite upconversion of entangled photons has been performed experimentally (*15*) [for a more detailed discussion of detection issues see (*11*)]. Does the enhancement persist at higher noise temperatures and for larger numbers of photons in the signal? What are the maximum enhancements obtainable via quantum illumination over multiphoton input states, including Gaussian states? These questions and many others must be answered before quantum illumination can prove itself useful in practice.

References and Notes

1. M. F. Sacchi, *Phys. Rev. A* **71**, 062340 (2005).
2. M. F. Sacchi, *Phys. Rev. A* **72**, 014305 (2005).
3. G. M. D'Ariano, M. F. Sacchi, J. Kahn, *Phys. Rev. A* **72**, 052302 (2005).
4. S.-K. Choi, M. Vasilyev, P. Kumar, *Phys. Rev. Lett.* **83**, 1938 (1999).
5. A. Gatti, E. Brambilla, L. A. Lugiato, M. Kolobov, *J. Opt. B Quant. Semiclassical Opt.* **2**, 196 (2000).
6. A. Mosset, F. Devaux, E. Lantz, *Phys. Rev. Lett.* **94**, 223603 (2005).
7. V. Giovannetti, S. Lloyd, L. Maccone, *Science* **306**, 1330 (2004).
8. M. O. Scully, M. S. Zubairy, *Quantum Optics* (Cambridge Univ. Press, Cambridge, 1997).
9. K. M. R. Audenaert *et al.*, *Phys. Rev. Lett.* **98**, 160501 (2007).
10. J. Calsamiglia, R. Munoz-Tapia, Ll. Masanes, A. Acin, E. Bagan, *Phys. Rev. A* **77**, 032311 (2008).
11. Materials and methods are available as supporting material on *Science* Online.
12. C. H. Bennett, S. J. Wiesner, *Phys. Rev. Lett.* **69**, 2881 (1992).
13. C. H. Bennett, P. W. Shor, J. A. Smolin, A. V. Thapliyal, *IEEE Trans. Inf. Theory* **48**, 2637 (2002).
14. H.-B. Fei, B. M. Jost, S. Popescu, B. E. A. Saleh, M. C. Teich, *Phys. Rev. Lett.* **78**, 1679 (1997).
15. B. Dayan, A. Pe'er, A. A. Friesem, Y. Silberberg, *Phys. Rev. Lett.* **93**, 023005 (2004).

16. This work was supported by the W. M. Keck foundation, Hewlett Packard, and Defense Advanced Research Projects Agency. The author thanks B. Erkmen, V. Giovannetti, S. Guha, L. Maccone, S. Pirandola, J. Shapiro, S.-H. Tan, M. Tsang, and H. Yuen for useful discussions.

Supporting Online Material

www.sciencemag.org/cgi/content/full/321/5895/1463/DC1
Materials and Methods

19 May 2008; accepted 1 August 2008
10.1126/science.1160627

reCAPTCHA: Human-Based Character Recognition via Web Security Measures

Luis von Ahn,* Benjamin Maurer, Colin McMillen, David Abraham, Manuel Blum

CAPTCHAs (Completely Automated Public Turing test to tell Computers and Humans Apart) are widespread security measures on the World Wide Web that prevent automated programs from abusing online services. They do so by asking humans to perform a task that computers cannot yet perform, such as deciphering distorted characters. Our research explored whether such human effort can be channeled into a useful purpose: helping to digitize old printed material by asking users to decipher scanned words from books that computerized optical character recognition failed to recognize. We showed that this method can transcribe text with a word accuracy exceeding 99%, matching the guarantee of professional human transcribers. Our apparatus is deployed in more than 40,000 Web sites and has transcribed over 440 million words.

A CAPTCHA (*1*, *2*) is a challenge-response test used on the World Wide Web to determine whether a user is a human or a computer. The acronym stands for Completely Automated Public Turing test to tell Computers and Humans Apart. A typical CAPTCHA is an image containing several distorted characters that appears at the bottom of Web registration forms. Users are asked to type the wavy characters to "prove" they are human. Current computer programs cannot read distorted text as well as humans can (*3*), so CAPTCHAs act as sentries against automated programs that attempt to abuse online services. Owing to their effectiveness as a security measure, CAPTCHAs are used to protect many types of Web sites, including free-e-mail providers, ticket sellers, social networks, wikis, and blogs. For example, CAPTCHAs prevent ticket scalpers from using computer programs to buy large numbers of concert tickets, only to resell them at an inflated price. Sites such as Gmail and Yahoo Mail use CAPTCHAs to stop spammers from obtaining millions of free e-mail accounts, which they would use to send spam e-mail.

According to our estimates, humans around the world type more than 100 million CAPTCHAs every day (see supporting online text), in each case spending a few seconds typing the distorted characters. In aggregate, this amounts to hundreds of thousands of human hours per day. We report on an experiment that attempts to make positive use of the time spent by humans solving CAPTCHAs. Although CAPTCHAs are effective at preventing large-scale abuse of online services, the mental effort each person spends solving them is otherwise wasted. This mental effort is invaluable, because deciphering CAPTCHAs requires people to perform a task that computers cannot.

We show how it is possible to use CAPTCHAs to help digitize typeset texts in nondigital form by enlisting humans to decipher the words that computers cannot recognize. Physical books and other texts written before the computer age are currently being digitized en masse (e.g., by the Google Books Project and the nonprofit Internet

Computer Science Department, Carnegie Mellon University, 5000 Forbes Avenue, Pittsburgh, PA 15213, USA.

*To whom correspondence should be addressed. E-mail: biglou@cs.cmu.edu

Archive) to preserve human knowledge and to make information more accessible to the world. The pages are photographically scanned and the resulting bitmap images are transformed into text files by optical character recognition (OCR) software. This transformation into text is useful because the books can then be indexed, searched, and stored in a format that can be easily analyzed and manipulated. One of the stumbling blocks in the digitization process is that OCR is far from perfect at deciphering the words in bitmap images of scanned texts. As we show below, for older prints with faded ink and yellowed pages, OCR cannot recognize about 20% of the words. By contrast, humans are more accurate at transcribing such print. For example, two humans using the "key and verify" technique, where each types the text independently and then any discrepancies are identified, can achieve more than 99% accuracy at the word level (*4*, *5*). Unfortunately, human transcribers are expensive, so only documents of extreme importance are manually transcribed.

Our apparatus, called "reCAPTCHA," is used by more than 40,000 Web sites (*6*) and demonstrates that old print material can be transcribed, word by word, by having people solve CAPTCHAs throughout the World Wide Web. Whereas standard CAPTCHAs display images of random characters rendered by a computer, reCAPTCHA displays words taken from scanned texts. The solutions entered by humans are used to improve the digitization process. To increase efficiency and security, only the words that automated OCR programs cannot recognize are sent to humans. However, to meet the goal of a CAPTCHA (differentiating between humans and computers), the system needs to be able to verify the user's answer. To do this, reCAPTCHA gives the user two words, the one for which the answer is not known and a second "control" word for which the answer is known. If users correctly type the control word, the system assumes they are human and gains confidence that they also typed the other word correctly (Fig. 1). We describe the exact process below.

We start with an image of a scanned page. Two different OCR programs analyze the image; their respective outputs are then aligned with each other by standard string matching algorithms (*7*) and compared to each other and to an English dictionary. Any word that is deciphered differently by both OCR programs or that is not in the English dictionary is marked as "suspicious." These are typically the words that the OCR programs failed to decipher correctly. According to our analysis, about 96% of these suspicious words are recognized incorrectly by at least one of the OCR programs; conversely, 99.74% of the words not marked as suspicious are deciphered correctly by both programs. Each suspicious word is then placed in an image along with another word for which the answer is already known, the two words are distorted further to ensure that automated programs cannot decipher them, and the resulting image is used as a CAPTCHA. Users are asked to type both words correctly before being allowed through. We refer to the word whose answer is already known as the "control word" and to the new word as the "unknown word." Each reCAPTCHA challenge, then, has an unknown word and a control word, presented in random order. To lower the probability of automated programs randomly guessing the correct answer, the control words are normalized in frequency; for example, the more common word "today" and the less common word "abridged" have the same probability of being served. The vocabulary of control words contains more than 100,000 items, so a program that randomly guesses a word would only succeed 1/100,000 of the time (*8*). Additionally, only words that both OCR programs failed to recognize are used as control words. Thus, any program that can recognize these words with nonnegligible probability would represent an improvement over state-of-the-art OCR programs.

To account for human error in the digitization process, reCAPTCHA sends every suspicious word to multiple users, each time with a different random distortion. At first, it is displayed as an unknown word. If a user enters the correct answer to the associated control word, the user's other answer is recorded as a plausible guess for the unknown word. If the first three human guesses match each other, but differ from both of the OCRs' guesses, then (and only then) the word becomes a control word in other challenges. In case of discrepancies among human answers, reCAPTCHA sends the word to more humans as an "unknown word" and picks the answer with the highest number of "votes," where each human answer counts as one vote and each OCR guess counts as one half of a vote (recall that these words all have been previously processed by OCR). In practice, these weights seem to yield the best results, though our accuracy is not very sensitive to them (as long as more weight is given to human guesses than OCR guesses). A guess must obtain at least 2.5 votes before it is chosen as the correct spelling of the word for the digitization process. Hence, if the first two human guesses match each other and one of the OCRs, they are considered a correct answer; if the first three guesses match each other but do not match either of the OCRs, they are considered a correct answer, and the word becomes a control word. To account for words that are unreadable, reCAPTCHA has a button that allows users to request a new pair of words. When six users reject a word before any correct spelling is chosen, the word is discarded as unreadable. After all suspicious words in a text have been deciphered, we apply a post-processing step because human users make a variety of predictable mistakes (see supporting online text). From analysis of our data, 67.87% of the words required only two human responses to be considered correct, 17.86% required three, 7.10% required four, 3.11% required five, and only 4.06% required six or more (this includes words discarded as unreadable).

A large-scale deployment of the system has enabled us to collect a number of findings (see supporting online text for more details about the deployment). The first finding is that the process of deciphering words with CAPTCHAs can match the highest-quality guarantee given by dedicated human transcription services. A random sample of 50 scanned articles from five different years (1860, 1865, 1908, 1935, and 1970) of the *New York Times* archive (http://nytimes.com) was chosen and manually transcribed by two professionals to estimate the per-word accuracy of reCAPTCHA, including the postprocessing corrections mentioned above. The total number of words was 24,080. Each word counted as a "hit" if the algorithm deciphered the entire word correctly or a "miss" if any of the letters were wrong. The error rate was defined as the number of misses divided by the

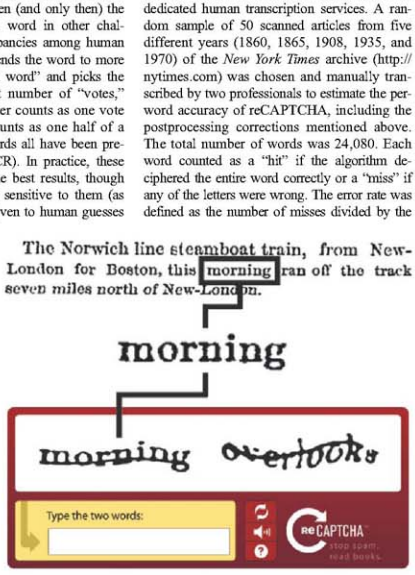


Fig. 1. The reCAPTCHA system displays words from scanned texts to humans on the World Wide Web. In this example, the word "morning" was unrecognizable by OCR. reCAPTCHA isolated the word, distorted it using random transformations including adding a line through it, and then presented it as a challenge to a user. Because the original word ("morning") was not recognized by OCR, another word for which the answer was known ("overlooks") was also presented to determine if the user entered the correct answer.

total number of words. To account for potential errors in the ground truth transcriptions, every miss was manually inspected in the ground truth and fixed in case it was an error. The results of one OCR program were run through the same process for comparison.

The reCAPTCHA system achieved an accuracy of 99.1% at the word level (216 errors out of 24,080 words), whereas the accuracy of standard OCR was only 83.5% (3976 errors). The percentage of words on which both OCR systems made a mistake was 7.3%. An accuracy of 99.1% is within the acceptable "over 99%" industry standard guarantee for "key and verify" transcription techniques in which two professional human transcribers independently type the data and discrepancies are corrected [e.g., see (*4*, *5*)]. As an anecdote, the professional manual transcriptions of the articles that were collected as "ground truth" to measure the accuracy of reCAPTCHA originally contained 189 errors, almost as many as those made by reCAPTCHA. There were many reasons for the mistakes made by reCAPTCHA, but the most common were issues of alignment and segmentation of the words by the OCR systems; for example, in some cases both OCR programs entirely missed a word or a set of words. The exact reason for errors in the professional human transcription is unknown to us, but the errors probably occur when the transcribers type a word differently from each other, and then a mistake is made in the correction of the discrepancy. The fact that reCAPTCHA can achieve an accuracy comparable to the "gold standard" accuracy of two independent humans can be counterintuitive because human transcribers make use of context (words immediately before and after), whereas words presented by reCAPTCHA are shown individually, in isolation from the original context. On the other hand, reCAPTCHA uses a combination of OCR and multiple humans, which in some cases turns out to be more resilient to accidental typographical mistakes.

Another finding is that CAPTCHAs constitute a viable mechanism to harness large amounts of human mental effort. After exactly 1 year of running the system, humans had solved more than 1.2 billion CAPTCHAs, amounting to over 440 million suspicious words correctly deciphered. Assuming 100,000 words per book (400 pages, 250 words per page), this is equivalent to over 17,600 books manually transcribed (about 25% of the words in each book are marked as suspicious by our algorithm). The system continues to grow in popularity: The rate of transcription currently exceeds 4 million suspicious words per day, which is equivalent to about 160 books per day. Achieving this rate via conventional "key and verify" means (without aid from OCR, so every word in a text would be typed) would require a workforce of more than 1500 people deciphering words 40 hours per week (assuming an average rate of 60 words per minute).

There are many reasons why Web sites choose to use reCAPTCHA. First, because we use only words from scanned books upon which OCR failed, reCAPTCHA is currently more secure than the conventional CAPTCHAs that generate their own randomly distorted characters. As shown by (*9*–*11*), it is possible to build algorithms that can read the distorted text generated by many widely used conventional CAPTCHAs with a success rate of more than 90% in some cases. As implemented by us, the same algorithms fail to recognize reCAPTCHA challenges 100% of the time. One reason for this is that the artificial distortions of characters in conventional CAPTCHAs come from a limited (and usually simple) distribution of possible transformations that remain readable to humans. Therefore, it is feasible to build machine learning algorithms that, after some training, can recognize the distorted characters. The words displayed by reCAPTCHA, however, have three types of distortions. First, and most importantly, there are natural distortions that result from the underlying texts having faded through time. Second, the scanning process introduces noise. Third, we introduce artificial transformations similar to those used by standard CAPTCHAs so the challenges remain difficult for computer programs even if there is an OCR slightly better than ours. Although there has been work in the image-processing community on modeling the natural degradation process of scanned books (*12*), such models are imperfect, so the distribution of distortions in reCAPTCHA is less limited. Additionally, reCAPTCHA displays only words on which two OCR programs failed. Because the words used as control words for reCAPTCHA constitute ~4% of the words from our distribution of scanned books, any program that can recognize a fraction p of the reCAPTCHA challenges without having compromised our database can be directly used to improve the accuracy of these OCR programs by a fraction of $0.04p$ on this same distribution (*13*). This would represent an advance in state-of-the-art OCR technology.

In essence, the words used by reCAPTCHA are the "hardest" words from scanned texts for computers to decipher. Humans, by contrast, can decipher the reCAPTCHA challenges with ease: Users of the Web sites that switch to using reCAPTCHA typically complain less often than when the sites used a different type of CAPTCHA. This is partly due to some users being more willing to accept reCAPTCHA because their work is contributing to the digitization of human knowledge [as can be seen by the vast number of blog entries praising reCAPTCHA (*14*)]. Additionally, based on more than 1 billion responses, the overall success rate for reCAPTCHA is 96.1%, a healthy number considering that a simple typing mistake would imply a failure. We do note that the success rate is not the same across all users. For example, non-English speakers seem to perform slightly worse than English speakers: Internet Protocol (IP) addresses from countries where the native language is not English have success rates that vary from 92.6 to 96.9%, depending on the country, whereas IP addresses from English-speaking countries range from 97.1 to 97.4% (because we have millions of data points, all of these differences are statistically significant with $P < 0.01$). Furthermore, even including English-speaking countries, conditioned on failing the first challenge, IP addresses that attempted a second challenge within 30 s have a success rate of only 89.9%. Another observation is that the success rate is proportional to the length of the control word: Four-character words have a success rate of 93.7%; five-character words, 95.7%; six-character words, 96.4%; seven-character words, 96.7%; etc. This can be explained by longer words providing more context for the users. The same relation holds when restricting attention to countries where the native language is not English, but to a lesser extent (consistent with our explanation that knowledge of the English language helps with longer words).

A second reason why Web sites adopt reCAPTCHA is that, although reCAPTCHA presents two words instead of just one, it typically takes no more time for users to solve a reCAPTCHA than to solve a standard CAPTCHA. Standard CAPTCHAs present six to eight randomly chosen characters (not an English word), which take about the same time to decipher as two English words. User testing on our site (http://captcha.net) showed that it took 13.51 s on average (SD = 6.37) for 1000 randomly chosen users to solve a seven-letter conventional CAPTCHA (25th percentile was 8.28 s, median was 12.62 s, and 75th percentile was 17.12 s), whereas it took 13.06 s on average (SD = 7.67) for a different set of 1000 randomly chosen users (also from http://captcha.net) to solve a reCAPTCHA (25th percentile was 5.79 s, median was 12.64 s, and 75th percentile was 18.91 s). The difference is not statistically significant, and indeed, the average time spent on reCAPTCHA was lower (although the median was 0.02 s higher). The fact that both standard CAPTCHAs and reCAPTCHAs take roughly the same amount of time to solve should not be surprising, because English words have patterns to which human users are accustomed. In addition, the time taken to solve reCAPTCHAs varies more widely because English words vary in length (*15*).

We believe the results presented here are part of a proof of concept of a more general idea: "Wasted" human processing power can be harnessed to solve problems that computers cannot yet solve. Some have referred to this idea as "human computation." In previous work (*16*–*18*), we have shown that such processing power can be harnessed through computer games: People play these games and, as a result, collectively perform tasks that computers cannot yet perform. Inspired by this work, biologists have recently built Fold It (http://fold.it/) (*19*), a game in which people compete to determine the ideal structure of a given protein. Here, we have shown that CAPTCHAs

constitute another avenue for "reusing" wasted computational power, while serving the useful purpose of preventing automated abuse over the Internet. A related, but different, line of work is ASIRRA (*20*), which has shown that CAPTCHAs can be used for humanitarian purposes. In this system, pictures of cats and dogs are presented to the user, who has to determine which ones are cats and which ones are dogs. The humanitarian twist is that the pictures come from animal shelters: If users like one of the cats or dogs, they can adopt it. More generally, computers do not perform as well as humans in visual recognition tasks. Perhaps a method similar to reCAPTCHA can be used to annotate or tag large quantities of images.

We hope that reCAPTCHA continues to have a positive impact on modern society by helping to digitize human knowledge.

References and Notes

1. L. von Ahn, M. Blum, N. Hopper, J. Langford, in *Advances in Cryptology*, E. Biham, Ed., vol. 2656 of *Lecture Notes in Computer Science* (Springer, Berlin, 2003), pp. 294–311.
2. L. von Ahn, M. Blum, J. Langford, *Commun. ACM* **47**, 56 (2004).
3. K. Chellapilla, K. Larson, P. Simard, M. Czerwinski, in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Association for Computing Machinery, New York, 2005), pp. 711–720.
4. F. A. Long, *Am. J. Econ. Sociol.* **52**, 223 (1993).
5. Advantage Research Incorporated, *Data Collection Services Quality Control* (www.advantageresearchinc.com/dcs_qualitycontrol.htm).
6. Based on the number of unique Web sites that have registered to use the reCAPTCHA service on www.recaptcha.net as of 15 July 2008.
7. This alignment and the word segmentation done by the OCR programs are imperfect and are the biggest sources of errors in our system.
8. Because computer programs can easily attempt to pass the CAPTCHA multiple times, if a computer has a success rate of even 5%, the CAPTCHA is considered broken. A typical convention is that a program should not be able to pass the CAPTCHA with a success rate of more than 1 in 10,000. (Downloading 10,000 CAPTCHA images requires substantial usage of bandwidth, exposing the IP address as potentially abusive.) Our system uses more than 100,000 words, which yields a probability of random guessing that is much smaller than 1/10,000. By contrast, conventional CAPTCHAs that use seven random characters yield an even smaller probability of success for random guessing: $1/36^7$.
9. K. Chellapilla, P. Y. Simard, in *Advances in Neural Information Processing Systems 17*, L. K. Saul, Y. Weiss, L. Bottou, Eds. (MIT Press, Cambridge, MA, 2005), pp. 265–272.
10. G. Mori, J. Malik, in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* **1**, 134 (IEEE Computer Society, Los Alamitos, CA, 2003).
11. A. Thayananthan, B. Stenger, P. H. S. Torr, R. Cipolla, in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* **1**, 127 (IEEE Computer Society, Los Alamitos, CA, 2003).
12. T. Kanungo, Q. Zheng, *IEEE Trans. Pattern Anal. Mach. Intell.* **26**, 520 (2004).
13. The assumption that the adversary's program has not compromised our database is necessary. The practical concern is not a hacker somehow infiltrating our system, but an adversary that poisons our database by submitting large quantities of bogus answers, because answers to the control words in reCAPTCHA come from the users themselves. Such an attack, however, is infeasible. For an unknown word to become a control word, the first three user answers must match each other and must have correct answers to the three different associated control words. An attacker randomly guessing the answer to three different control words would have a probability of success of $1/10^{15}$.
14. Based on manual review of the results from a search of http://blogsearch.google.com/ for the term "recaptcha" on 21 July 2008.
15. These timing numbers are for users whose IP addresses come from both English- and non-English–speaking countries. Restricting the timing numbers to IP addresses from non-English–speaking countries yields an average of 16.32 s for reCAPTCHA and 13.91 s for the standard CAPTCHA.
16. L. von Ahn, *Computer* **39**, 92 (2006).
17. L. von Ahn, L. Dabbish, in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Association for Computing Machinery, New York, 2004), pp. 319–326.
18. L. von Ahn, R. Liu, M. Blum, in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Association for Computing Machinery, New York, 2004), pp. 55–64.
19. *Economist* **387**, 105 (2008).
20. J. Elson, J. Douceur, J. Howell, in *Proceedings of the 14th ACM Conference on Computer and Communications Security* (Association for Computing Machinery, New York, 2007), pp. 366–374.

21. This work was partially supported by gifts from the Heinz Endowment and the Fine Foundation, by an equipment grant from Intel Corporation and a research grant from the New York Times Company, and by the Army Research Office through grant number DAAD19-02-1-0389 to CyLab at Carnegie Mellon University. L.v.A. was partially supported by a Microsoft Research New Faculty Fellowship and a MacArthur Fellowship. L.v.A. is cofounder, president, and chief executive officer of reCAPTCHA, a free CAPTCHA service that helps to digitize books. B.M., C.M., D.A., and M.B. are also cofounders of reCAPTCHA. L.v.A., B.M., and M.B. have filed a patent related to the technology described in this paper.

Supporting Online Material

www.sciencemag.org/cgi/content/full/1160379/DC1
SOM Text

12 May 2008; accepted 5 August 2008
Published online 14 August 2008;
10.1126/science.1160379
Include this information when citing this paper.

A Rubberlike Stretchable Active Matrix Using Elastic Conductors

Tsuyoshi Sekitani,[1] Yoshiaki Noguchi,[1] Kenji Hata,[2] Takanori Fukushima,[3,4]* Takuzo Aida,[3,4] Takao Someya[1,5]†

By using an ionic liquid of 1-butyl-3-methylimidazolium bis(trifluoromethanesulfonyl)imide, we uniformly dispersed single-walled carbon nanotubes (SWNTs) as chemically stable dopants in a vinylidene fluoride-hexafluoropropylene copolymer matrix to form a composite film. We found that the SWNT content can be increased up to 20 weight percent without reducing the mechanical flexibility or softness of the copolymer. The SWNT composite film was coated with dimethylsiloxane–based rubber, which exhibited a conductivity of 57 siemens per centimeter and a stretchability of 134%. Further, the elastic conductor was integrated with printed organic transistors to fabricate a rubberlike active matrix with an effective area of 20 by 20 square centimeters. The active matrix sheet can be uniaxially and biaxially stretched by 70% without mechanical or electrical damage. The elastic conductor allows for the construction of electronic integrated circuits, which can be mounted anywhere, including arbitrary curved surfaces and movable parts, such as the joints of a robot's arm.

The creation of stretchable electronics is one of the most interesting challenges in materials science and engineering. Stretchability is an entirely different concept from the miniaturization trend pursued by conventional electronics, and thus has the potential to open exciting opportunities, particularly in the area of large-area electronics (*1*–*6*). In the past decade, large-area electronic devices have become thin and light enough to allow the fabrication of large-area solar cells (*7*) and displays easily hung on roofs and walls. It is expected that large-area electronic devices will now be developed further, making the realization of bendable and rollable displays possible (*8*). At the same time, large-area flexible sensors (*9*, *10*) and actuators (*11*, *12*) are another emerging frontier. Although these achievements represent valuable advances, the utility of flexible electronics is limited to nearly flat substrates such as walls or paper. In contrast, stretchable electronics can cover arbitrary curved surfaces and movable parts such as the joints of a robot's arm, and thus would substantially expand where electronics can be used.

Efforts toward stretchable electronics have originated from electronics using metal electrodes on rubber substrates (*13*–*16*). Khang *et al.*, Sun *et al.*, and Kim *et al.* have embedded active

[1]Quantum-Phase Electronics Center, School of Engineering, University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8656, Japan. [2]Research Center for Advanced Carbon Materials, National Institute of Advanced Industrial Science and Technology, Tsukuba 305-8565, Japan. [3]Department of Chemistry and Biotechnology, School of Engineering, University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8656, Japan. [4]Nanospace Project, Exploratory Research for Advanced Technology–Solution Oriented Research for Science and Technology, Japan Science and Technology Agency, National Museum of Emerging Science and Innovation, 2-41 Aomi, Koto-ku, Tokyo 135-0064, Japan. [5]Collaborative Institute for Nano Quantum Information Electronics, University of Tokyo, 4-6-1, Komaba, Meguro-ku, Tokyo 153-8505, Japan.

*Present address: Advanced Science Institute, RIKEN, 2-1 Hirosawa, Wako, Saitama 351-0198, Japan.
†To whom correspondence should be addressed. E-mail: someya@ap.t.u-tokyo.ac.jp

components, such as transistors and diodes, in rubber sheets and integrated them with wavy metal wires by carefully controlling the strains in thin films (*14*–*16*). Their electrical circuits have high mechanical durability and show good electrical performance under stretching because all the circuit components are stretchable. In an alternative cost-effective approach, integrated circuits have also been directly fabricated on plastic films and mechanically processed to form perforated films with net-shaped structures, which serve as stretchable artificial skins (*10*). Net-shaped integrated circuits are flexible but have inelastic wirings; their mechanical robustness can be substantially improved by using elastic wirings.

The simultaneous incorporation of excellent mechanical robustness and electronic performance is the key to realizing rubberlike stretchable electronics. Rigid materials usually exhibit good electronic performance and controllability or stability, but they exhibit poor mechanical robustness. On the other hand, soft materials show good mechanical properties while exhibiting poor electronic properties. For example, a conductivity of 0.1 S/cm for the best elastomers (rubbers filled with carbon particles) is insufficient to operate integrated circuits.

Besides metals and carbon-particle composites, carbon-nanotube–based conducting materials have also been produced (*17*–*20*). Their conductivities are typically 10^{-3} to 10^{1} S/cm (*17*–*19*). Furthermore, carbon-nanotube and/or conducting-polymer (polyaniline) composites have been reported (*20*); however, these materials cannot be stretched.

We developed rubberlike transistor-active matrices that can be stretched biaxially by 70%. The key was the fabrication of a highly elastic conductor and gel with high conductivity composed from millimeter-long single-walled carbon nanotubes (SWNTs) (*21*, *22*), an ionic liquid, and a compatible fluorinated copolymer (*23*). We used SWNTs as a conducting dopant because they are chemically inert and can improve the mechanical properties of polymer matrices. It has been reported that fine bundles of SWNTs can be produced by grinding with imidazolium ion-based ionic liquids (*24*, *25*).

SWNT composite films (referred to as SWNT films henceforth) were fabricated using millimeter-long SWNTs, an ionic liquid, and a fluorinated copolymer. A schematic representation of the fabrication process of the elastic conductors is shown in Fig. 1. We used super-growth SWNTs (>99.98% in purity, >1 mm in length, and 3 nm in diameter) (*22*) unless otherwise specified. Typically, we mixed the SWNTs (50 mg) with 50 mg of an ionic liquid [1-butyl-3-methylimidazolium bis(trifluoromethanesulfonyl)imide (BMITFSI)] and subjected the resulting suspension to an automatic grinding system for 1 hour, giving rise to a black pastelike substance we refer to as "bucky gel." To the gel (100 mg) we successively added 4-methly-2-pentanone (8 ml) and a fluorinated copolymer, vinylidene fluoride-hexafluoropropylene copolymer [100 mg; Daiel-G801 (Daikin, Osaka, Japan), weigh average molecular weight (*Mw*) of 150,000; referred to as G801 henceforth]. The mixture was stirred at 25°C (1 hour) and sonicated [UH-50 (SMT, Tokyo, Japan)] at 30°C (1 hour). After stirring again at 80°C (1 hour), the resulting swollen gel was poured onto a glass plate by drop casting and then air-dried for 24 hours to obtain a composite film (SWNT film), as shown in Fig. 2, A and B. With this method, we dispersed SWNTs uniformly into a fluorinated copolymer matrix to fabricate a SWNT film.

The advantage of fabrication using ionic liquids is that after the dispersion of SWNTs, ionic liquids can be recovered quantitatively (99%) by Soxhlet extraction and recycled for the next batch process. After the extraction of ionic liquids, conductivity decreases; however, it is as high as 10 S/cm, as shown later. In the following experiments, we used elastic conductors with ionic liquids.

The SWNT film is flexible and tensile, but it has low elasticity. In order to improve its tensibility and elasticity, it was mechanically processed with a numerically controlled punching system and transformed into a perforated film with a net-shaped structure (fig. S1). Subsequently, it was coated with dimethyl-siloxane–based silicone rubbers [polydimethylsiloxane (PDMS), Sylgard 184 or SH9555 (Dow Corning, Midland, MI)]. The resulting composite material is referred to as an "elastic conductor." Its elasticity is determined by

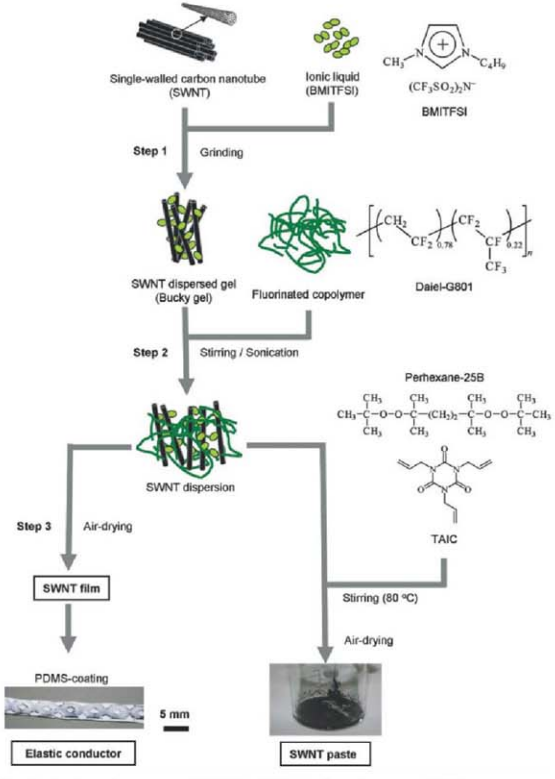


Fig. 1. Manufacturing process of SWNT film, SWNT elastic conductor, and SWNT paste.

the elasticity of PDMS, which is 6 N/mm^2. PDMS is nonadherent to almost all materials; however, our SWNT film has an excellent capillary surface (meaning a significantly large surface area), the same as SWNTs inherently possess. Therefore, by dip-coating diluted PDMS around SWNT films, PDMS was attached firmly to SWNT material.

We investigated the electrical and mechanical properties of the SWNT film and elastic conductor under tensile stress. The stretch test was performed with a high-precision mechanical system [Autograph/AG-X (Shimadzu, Kyoto, Japan)], and conductivity was measured by means of the four-probe method with a high-precision multimeter [34401A (Agilent, Santa Clara, CA)]. Figure 2C shows conductivity as a function of uniaxial tensile strain. For comparison, the stretchability and conductivity of a commercial conducting rubber containing carbon particles (Kinugawa Rubber Industrial, Chiba, Japan) are also shown in Fig. 2C. Although the stretchability of the commercial conducting rubber exceeded 150%, its conductivity was approximately 0.1 S/cm for all strains, which is insufficient for electronic circuit applications. In contrast, the SWNT film exhibited an extraordinarily high conductivity of 57 S/cm, and it did not show significant changes in conductivity or mechanical damage when uniaxially stretched by 38% or less. Furthermore, the SWNT elastic conductor exhibited a high conductivity of 57 S/cm, and it could be uniaxially stretched up to 134%; however, the conductivity decreased moderately with the application of tensile strain. Even under such a large strain, the material showed good conductivity (6 S/cm). We observed that the shapes of holes of elastic conductors changed during stretching. The enhancement in the stretchability of the net-shaped structure was realized because the interconnecting struts buckled and twisted out of plane upon tension (fig. S1).

To investigate the reversibility of the elastic conductor, several stretching cycles were applied. Measurements were performed at each stretching cycle by using the four-probe method. Figure 2D shows the normalized conductance as a function of the number of uniaxial stretching cycles. There was no significant change in conductance even after 4000 25%-stretching cycles, 500 50%-stretching cycles, 20 to 50 70%-stretching cycles, or 1 to 2 110%-stretching cycles. Further increase in strain beyond 110% caused an irreversible change in conductance, although it was greater than 1 S/cm. Similar to other stretchable materials,

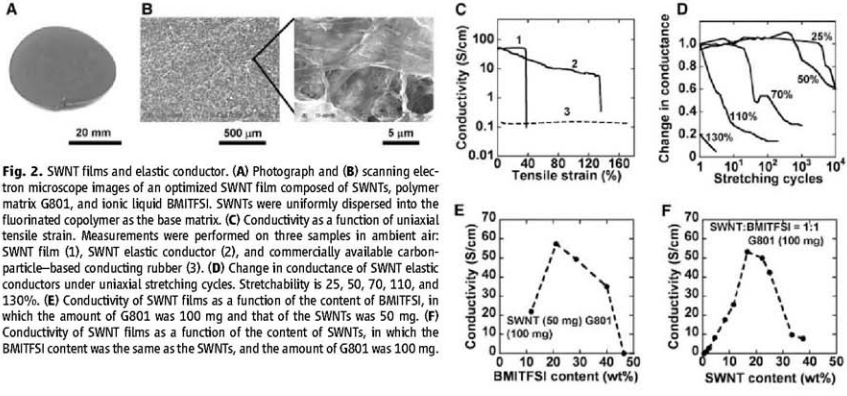


Fig. 2. SWNT films and elastic conductor. **(A)** Photograph and **(B)** scanning electron microscope images of an optimized SWNT film composed of SWNTs, polymer matrix G801, and ionic liquid BMITFSI. SWNTs were uniformly dispersed into the fluorinated copolymer as the base matrix. **(C)** Conductivity as a function of uniaxial tensile strain. Measurements were performed on three samples in ambient air: SWNT film (1), SWNT elastic conductor (2), and commercially available carbon-particle–based conducting rubber (3). **(D)** Change in conductance of SWNT elastic conductors under uniaxial stretching cycles. Stretchability is 25, 50, 70, 110, and 130%. **(E)** Conductivity of SWNT films as a function of the content of BMITFSI, in which the amount of G801 was 100 mg and that of the SWNTs was 50 mg. **(F)** Conductivity of SWNT films as a function of the content of SWNTs, in which the BMITFSI content was the same as the SWNTs, and the amount of G801 was 100 mg.

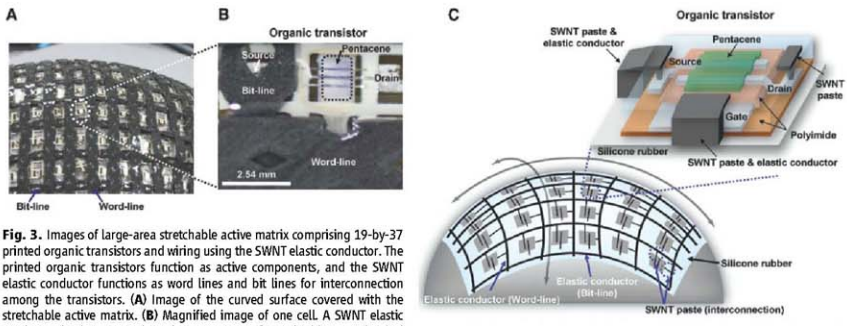


Fig. 3. Images of large-area stretchable active matrix comprising 19-by-37 printed organic transistors and wiring using the SWNT elastic conductor. The printed organic transistors function as active components, and the SWNT elastic conductor functions as word lines and bit lines for interconnection among the transistors. **(A)** Image of the curved surface covered with the stretchable active matrix. **(B)** Magnified image of one cell. A SWNT elastic conductor having a net-shaped structure was formed with a mechanical punching system and then coated with silicone rubbers. **(C)** Schematic illustrations of the stretchable active matrix.

either conducting or nonconducting, our elastic conductor exhibits irreversible mechanical and electrical changes after many stretching cycles (Fig. 2D). However, our conductor with a conductivity of 50 S/cm can be stretched approximately 500 times at a high level of strain (50%) without notable degradation. The feasibility of such a robust conductor has been demonstrated unambiguously by using the net-shaped structure with a PDMS coating. In our study, the electrical and mechanical characteristics of the conductor did not deteriorate over time—not over at least 1 year—because its components were chemically stable.

The fabrication process is fully optimized to achieve the best mechanical and conductive properties of the SWNT composite film. To establish this, we examined various conditions by changing process parameters and materials. For instance, as a polymer matrix, we examined vinylidene fluoride-hexafluoropropylene copolymer with composition ratios of 0.78:0.22 (G801) and 0.88:0.12 [KYNAR-FLEX2801 (Arkema, Paris, France), molecular weight of 30,000; referred to as KYNAR henceforth]. For ionic liquids, we used, in addition to BMITFSI, 1-butyl-3-methylimidazolium hexafluorophosphate ($BMIPF_6$) and 1-butyl-3-methylimidazolium tetrafluoroborate ($BMIBF_4$) (fig. S2).

One of the most important experiments in this study was to determine a combination of elastic polymers and ionic liquids that are compatible with each other. When G801 and BMITFSI were used, the resulting SWNT films exhibited a very smooth, flat, and uniform surface (Fig. 2, A and B). In contrast, all other combinations tested in this study (fig. S2) did not exhibit good compatibility. For example, KYNAR, which is a hard resin, was not compatible with BMITFSI, and the SWNT film obtained with this combination was readily crumpled or deformed.

Besides the compatibility of the materials, the correct mixing ratios of SWNT, BMITFSI, and G801 are needed to fabricate mechanically robust and highly conductive SWNT films. First, the BMITFSI content was changed from 12 to 47 weight percent (wt %), and the amounts of SWNT and G801 were 50 and 100 mg, respectively (Fig. 2E and fig. S4). When the BMITFSI content was greater than 40 wt %, the SWNT film became discontinuous; when it was less than 10 wt %, the film was thick and porous, and thus was fragile and had low conductivity. We found that the highest conductivity can be obtained when the contents of SWNT and BMITFSI are both 20 wt %, without sacrificing the mechanical flexibility of the copolymer (Fig. 2F and figs. S5 and S6). Next, the bucky gel with optimized SWNT and BMITFSI contents was mixed with G801 by varying the mixture ratios between the bucky gel and G801 (fig. S6). Figure 2F displays the plots of the conductivity of the film as a function of the SWNT content. When the SWNT content was less than 10 wt %, the SWNT film was discontinuous (fig. S5). In conformance with the results shown in Fig. 2E, when the SWNT content was 16%, the conductivity was as high as 53 S/cm, indicating excellent reproducibility and controllability of our fabrication method for highly elastic conductors. Meanwhile, when the SWNT content was greater than 30 wt %, the resultant film was thick and porous, thereby giving rise to brittleness and low conductivity.

The extremely high aspect ratio of the super-growth SWNTs (*22*) is important not only for acquiring higher conductivity but also for the stability of conductivity under tensile strain, because the conductivity of the super-growth SWNTs does not change with strain (fig. S8). In sharp contrast, the electrical properties of films fabricated with high-pressure carbon monoxide process (HiPco) SWNTs showed a strong dependence on stretchability; their conductivity rapidly decreased with an increase in the applied tensile stress (fig. S8).

A previous study on circuits containing stretchable interconnects (*13*) revealed that the contact points between stretchable and rigid materials are locations where mechanical failures can occur. We fabricated a SWNT-based paste by using the G801-containing bucky gel via cross-linking the polymer matrix (Fig. 1). A peroxide cross-linking initiator [Perhexane-25B (NOF, Tokyo, Japan)] (1.3 mg) and a cross-linker Triallylisocyanurate (TAIC) [(Nippon Kasei Chemical, Tokyo, Japan)] (4 mg) were added to the SWNT-dispersed solution. The mixture was stirred at 80°C for 1 hour and air dried, producing a paste with a partially cross-linked fluorinated copolymer. The paste exhibited a high conductivity of 5 to 10 S/cm and high adhesion capability. The adhesive strength increased up to 1.5 N/mm^2 (fig. S9). Therefore, the SWNT paste can be used to form interconnections between the contact pads of organic transistors (*8–12*, *26–31*) and SWNT-based elastic conductors that require considerably more mechanical robustness than other components, such as wires (Fig. 3).

Taking full advantage of the SWNT elastic conductor and paste, we manufactured a 19-by-37–cell organic transistor-based stretchable active matrix (Fig. 3) by combining printing, vacuum evaporation, and mechanical processes [supporting online material (SOM) and figs. S10 to S12] (*23*). We fabricated an array of organic transistors that were connected with each other by elastic con-

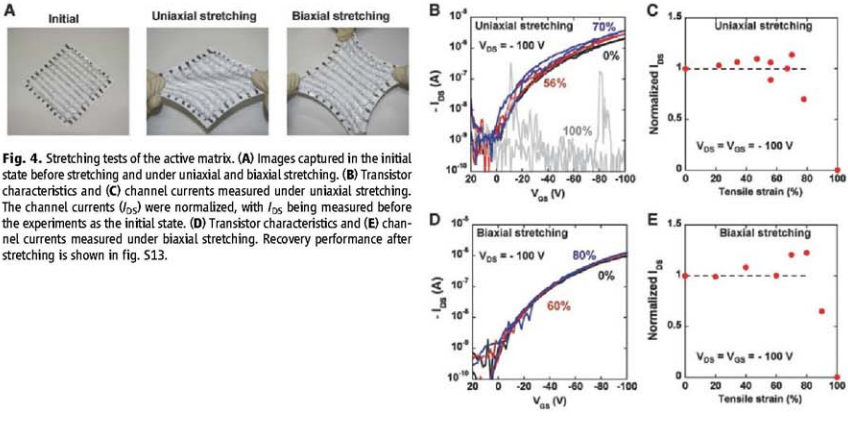


Fig. 4. Stretching tests of the active matrix. **(A)** Images captured in the initial state before stretching and under uniaxial and biaxial stretching. **(B)** Transistor characteristics and **(C)** channel currents measured under uniaxial stretching. The channel currents (I_{DS}) were normalized, with I_{DS} being measured before the experiments as the initial state. **(D)** Transistor characteristics and **(E)** channel currents measured under biaxial stretching. Recovery performance after stretching is shown in fig. S13.

ductors (bit lines for source electrodes or word lines for gate electrodes). The transistor characteristics did not change after the formation of the interconnections with the fabricated elastic conductors (fig. S11). The SWNT elastic conductors were connected to the contact pads for gate, source, and drain electrodes by the SWNT paste.

The sheet was stretched by increasing the tensile strain, and the transistor characteristics on the stretched sheet were examined (Fig. 4). In order to apply accurate strains uniformly to the transistor active matrix, we used a small active matrix comprising 10 by 10 cells. Furthermore, we measured the characteristics of the transistor positioned at the center of the matrix, because that transistor is stretched to a greater extent than those at the corners. When the sheet was stretched by 70% or less, the change in the transistor characteristics was negligibly small. Subsequently, the extension distortion was released, and recovery tests for the measurement of the transistor characteristics without strain were performed (fig. S13). After it was stretched by 70% or less, the transistor characteristics without strain did not change as compared with those measured before the stretching tests. Furthermore, no changes in electrical characteristics or mechanical damages were observed after 30 cycles at 70% stretching, indicating the excellent electrical functionalities and stability of the active matrix under stretching. However, stretching by 70% or more resulted in irreversible degradation. This is mainly due to the exfoliation of printed Ag electrodes from plastic substrates. The disconnection of organic transistors from the SWNT elastic conductor is quite rare because of the good adhesion of SWNT pastes.

The present approach in which semi-rigid active components are connected to each other through stretchable wirings is complementary to fully elastic circuits reported by Khang *et al.*, Sun *et al.*, and Kim *et al.* (*14*–*16*). Their circuits exhibit good mechanical durability under stretching; large deformations of stretchable active components may lead to changes in their electrical characteristics. In our approach, the electrical characteristics of active components do not change during stretching because they are not deformed. Therefore, the two approaches should be adopted according to the purpose of the application.

The conductivity of the SWNT-based elastic conductor is sufficiently high for application to high-performance and large-area electronic circuits. In the case of commercial conducting rubber using carbon particles (~0.1 S/cm), however, the resistance of electrical wiring between transistors (word and bit lines) is higher than the resistance of transistors in one state (fig. S14). This causes a decrease in channel currents, delay in circuits, and other detrimental effects.

Although we demonstrated the feasibility of SWNT-based elastic conductors and paste by fabricating an active matrix, the materials and integration technology can be applied to various types of electronic functionalities. When the stretchable active matrix is integrated with a two-dimensional array of pressure sensors, a rubber-like artificial skin can be obtained. Such sensor and actuator applications (*9*–*12*) do not always require components with dimensions of several tens or hundreds of micrometers. To realize this integration, it is necessary to develop mechanically robust, via interconnections. Furthermore, if the stretchable active matrix is integrated with an array of actuators and mounted on a curved surface, the touch feeling on the surface will be changed electrically. In this manner, the elastic conductor developed in this work enables electronic circuits to be mounted at locations where to date we have been unable to provide electrical functionalities. This is an important step toward producing intelligent surfaces as friendly human/electronics interfaces; in the future, such intelligent surfaces will be able to interact with people, objects, and the environment in new ways.

References and Notes

1. F. Garnier, R. Hajlaoui, A. Yassar, P. Srivastava, *Science* **265**, 1684 (1994).
2. J. A. Rogers, *Science* **291**, 1502 (2001).
3. C. D. Dimitrakopoulos, P. R. L. Malenfant, *Adv. Mater.* **14**, 99 (2002).
4. R. F. Service, *Science* **304**, 675 (2004).
5. H. E. Katz, *Chem. Mater.* **16**, 4748 (2004).
6. M. Hamedi, R. Forchheimer, O. Inganäs, *Nat. Mater.* **6**, 357 (2007).
7. D. B. Mitzi, L. L. Kosbar, C. E. Murray, M. Copel, A. Afzali, *Nature* **428**, 299 (2004).
8. J. A. Rogers *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **98**, 4835 (2001).
9. T. Someya *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **101**, 9966 (2004).
10. T. Someya *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 12321 (2005).
11. Y. Kato *et al.*, *IEEE Trans. Electron. Devices* **54**, 202 (2007).
12. T. Sekitani *et al.*, *Nat. Mater.* **6**, 413 (2007).
13. S. P. Lacour, J. Jones, S. Wagner, T. Li, Z. G. Suo, *Proc. IEEE* **93**, 1459 (2005).
14. D. Y. Khang, H. Q. Jiang, Y. Huang, J. A. Rogers, *Science* **311**, 208 (2006).
15. Y. G. Sun, W. M. Choi, H. Q. Jiang, Y. G. Y. Huang, J. A. Rogers, *Nat. Nanotechnol.* **1**, 201 (2006).
16. D. H. Kim *et al.*, *Science* **320**, 507 (2008).
17. R. Haggenmueller, H. H. Gommans, A. G. Rinzler, J. E. Fischer, K. I. Winey, *Chem. Phys. Lett.* **330**, 219 (2000).
18. C. Park *et al.*, *Chem. Phys. Lett.* **364**, 303 (2002).
19. B. S. Shim *et al.*, *Chem. Mater.* **19**, 5467 (2007).
20. G. M. Spinks, V. Mottaghitalab, M. B. Samani, P. G. Whitten, G. G. Wallace, *Adv. Mater.* **18**, 637 (2006).
21. S. Iijima, T. Ichihashi, *Nature* **363**, 603 (1993).
22. K. Hata *et al.*, *Science* **306**, 1362 (2004).
23. Materials and methods are available as supporting material on *Science* Online.
24. T. Fukushima *et al.*, *Science* **300**, 2072 (2003).
25. T. Fukushima, T. Aida, *Chem. Eur. J.* **13**, 5048 (2007).
26. Z. Bao, A. Dodabalapur, A. J. Lovinger, *Appl. Phys. Lett.* **69**, 4108 (1996).
27. A. Dodabalapur, L. Torsi, H. E. Katz, *Science* **268**, 270 (1995).
28. C. D. Dimitrakopoulos, S. Purushotham, J. Kymissis, A. Callegari, J. M. Shaw, *Science* **283**, 822 (1999).
29. H. Klauk, D. J. Gundlach, J. A. Nichols, T. N. Jackson, *IEEE Trans. Electron. Devices* **46**, 1258 (1999).
30. H. Klauk, U. Zschieschang, J. Pflaum, M. Halik, *Nature* **445**, 745 (2007).
31. N. Stutzmann, R. H. Friend, H. Sirringhaus, *Science* **299**, 1881 (2003).

32. This study was partially supported by the Grant-in-Aid for Scientific Research (Kakenhi, Wakate S), the Special Coordination Funds for Promoting and Technology. We thank T. Sakurai and M. Takamiya of University of Tokyo for invaluable discussion; Y. Komatsu of Daikin for fluorinated copolymer; and A. Kosaka of Exploratory Research for Advanced Technology–Solution Oriented Research for Science and Technology Nanospace Project, Japan Science and Technology Agency, for technical support. T. Someya and T. Sekitani have filed a provisional U.S. patent (61/064,024) on this work.

Supporting Online Material

www.sciencemag.org/cgi/content/full/1160309/DC1
Materials and Methods
Figs. S1 to S14
References

12 May 2008; accepted 21 July 2008
Published online 7 August 2008;
10.1126/science.1160309
Include this information when citing this paper.

Imaging of Transient Structures Using Nanosecond in Situ TEM

Judy S. Kim,[1,2]* Thomas LaGrange,[1] Bryan W. Reed,[1] Mitra L. Taheri,[1] Michael R. Armstrong,[1] Wayne E. King,[1] Nigel D. Browning,[1,2] Geoffrey H. Campbell[1]

The microstructure and properties of a material depend on dynamic processes such as defect motion, nucleation and growth, and phase transitions. Transmission electron microscopy (TEM) can spatially resolve these nanoscale phenomena but lacks the time resolution for direct observation. We used a photoemitted electron pulse to probe dynamic events with "snapshot" diffraction and imaging at 15-nanosecond resolution inside of a dynamic TEM. With the use of this capability, the moving reaction front of reactive nanolaminates is observed in situ. Time-resolved images and diffraction show a transient cellular morphology in a dynamically mixing, self-propagating reaction front, revealing brief phase separation during cooling, and thus provide insights into the mechanisms driving the self-propagating high-temperature synthesis.

Transmission electron microscopy (TEM) has evolved dramatically in recent years with the development of monochromation and spherical aberration correction (*1*–*3*) facilitating sub-angstrom spatial resolution. In situ TEM studies have also progressed from simple heating and cooling experiments to include capabilities such as nanoindentation (*4*) and environmental cells (*5*, *6*). The TEM is already a powerful tool for material characterization in diverse scien-

tific fields, but there is a need to incorporate fast time-resolution capabilities into EM.

In the past, direct electron imaging has been essential to uncover phenomena that are difficult to distinguish with diffraction (*7*–*9*), for instance, dislocation dynamics. Electron interrogation methods also have the potential to provide higher sensitivity and resolution compared with laser or x-ray methods. This is possible because of bright electron sources, the ability to control and focus electrons for different contrast mechanisms, and the stronger interaction of electrons with matter (*10*), resulting in a broad class of observable samples and length scales.

Traditional in situ TEM spatially resolves microstructural details of phase, structure, and morphology; however, it cannot collect data with acquisition times less than 1 ms. This is often orders of magnitude too slow to capture the essential material details of interface motion, crystal formation, twinning, and many other fundamental material processes. Many such processes are nonrecurring, necessitating single-shot techniques that capture images or diffraction patterns in a single brief exposure. At nanosecond time scales, such transient data are only attainable with the use of a single, short electron bunch with very high peak current by TEM standards (microamperes to milliamperes). Nanosecond-scale in situ TEM has been achieved using a conventional TEM modified to introduce a laser for stimulation of a photoemission electron source (*11*, *12*). Figure S1 illustrates how a 15-ns yttrium-lithium-fluoride–Nd laser frequency quintupled to 211 nm is directed toward the electron source, generating in excess of 2×10^9 photoemitted electrons in a single 15-ns bunch. Up to 5×10^7 of these electrons can be collimated into a small (micrometer-scale) area to probe the sample. This very high current density enables nanosecond single-shot imaging and electron diffraction. Imaging and diffraction provide complementary information, where diffraction reveals the evolution of crystallographic structure (*13*) (or amorphous nature) and imaging reveals details of the real-space evolution of fundamental nonrecurring processes in nanoscale materials dynamics.

The single-shot capability provides a pump-probe platform that can acquire a "snapshot" image or diffraction pattern of a transient process. The dynamic processes are stimulated in a TEM specimen by an additional pulsed laser (yttrium-aluminum-garnet–Nd), the pump laser. This laser heats or ablates a small area of the sample, initiating the process that will later be captured by the electron pulse. The time delay between the two pulses can be precisely controlled within a scale ranging from nanoseconds to hundreds of microseconds, with a timing jitter of ±1 ns. Longer multi-microsecond time delays used in this study cannot be achieved in a single laser–plus–beam splitter system without using optical paths of thousands of meters. By accumulating single-shot pump-probe observations with varying time delays, one can develop an understanding of each stage in the evolution of a material process.

Reactive multilayer foils (RMLFs), also termed nanostructured metastable intermolecular composites, are layers of reactant materials that undergo exothermic, self-propagating reactions when layer mixing is induced by an external energy source (Fig. 1). These nano foils exhibit mobile, high-temperature reaction zones where atoms of adjoining layers diffuse across the interfaces, wherein velocity and temperature can be manipulated by composition and geometry of the component materials (*14*–*16*). They are used as customized heat sources for rapid fuses, biological neutralization, and joining of materials by means of localized heating rather than global device heating (*17*, *18*).

The reactive foils contain stored chemical energy in the form of layered structures with <1 nm of interdiffusion at the interfaces (fig. S2). The reaction-front velocity (reaching ~10 m/s) is related to bilayer thickness, with an exception in thin bilayers (~10 nm) where intermixing during deposition retards the reaction propagation (*15*). Maximum temperatures attained during mixing vary greatly based on composition and geometry, reaching upwards of 1750 K (*19*). Little is known about the dynamic and transient events that transpire in the proximity of the reaction front, although these will govern the structure, mechanical properties, and performance of the reacted materials. The combination of nano-scale geometry and rapid velocity often make traditional characterization methods incapable of direct observation of the self-propagating high-temperature synthesis.

We use dynamic TEM (DTEM) to observe the RMLF reaction front for rapid phase-transition times and metastable morphologies. RMLF experimental samples are composed of five bilayers of $Al/Ni_{0.91}V_{0.09}$ (at a 2:3 atomic ratio, totaling 125 nm in thickness) and mounted in Cu mesh-clamping grids. Through fast optical imaging at 2 μs per frame (fig. S3 and movie S1), we found that when initiated by a 1064-nm laser pulse in an external vacuum chamber, the reaction front travels across the sample at a velocity of 13 m/s. Because the spatial resolution was 35 μm per pixel, the optical videos did not show additional definitive information. We then conducted DTEM experiments at the same pump-laser conditions [80-μm $1/e^2$ radius, 12-μJ energy, 3-ns full width at half maximum (FWHM) pulse duration]. The orientation of the experimental setup illustrated in the schematic (fig. S1, inset, and fig. S3) shows how the reaction front radiates out in all directions from the reaction initiation site in the freestanding RMLF membrane. By initiating the mixing reaction several hundred micrometers away from the observation area, the reaction front of the intermetallic formation zone is observed while it is in a steady state of propagation. Delays between the reaction-front arrival and the probe electron pulse were varied from ~0.7 to 25 μs, before the arrival and long after the front has passed.

At low magnification with a field of view of ~500 μm^2, the reaction front is visible after 1 μs as it travels away from the drive-laser initiation site (fig. S4). At increased magnification, the bright field images (Fig. 2) demonstrate the transient morphology of the reaction front. The reaction front is closely followed by the elongated finger-like structure, defining the interface between the reacted and unreacted material. These cellular features of Fig. 2 have a periodicity of ~600 nm wide and are greater than 40 μm long (equivalent to 3 μs of propagation time), showing a distinct formation and gradual termination, which proves their transient character.

The images in Fig. 2 are in plan view (with the viewing direction perpendicular to the metallic layers), and the cellular features are normal to the bilayer structure in the direction of the yellow arrow in the schematic Fig. 1. The size scale of the transient structures (600-nm periodicity) greatly exceeds the bilayer periodicity of the RMLF (25 nm); thus, the observed cellular structures are not directly related to the metallic-layer structure but instead suggest a transverse patterning mechanism inherent in the reaction process. This is reminiscent of binary alloy solidification in the midst of slight undercooling and also of solid solution eutectoid transformation (*20*, *21*). These models have shown that similar

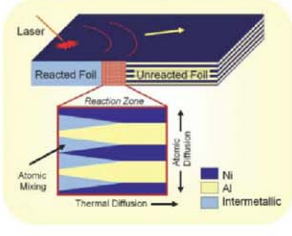


Fig. 1. Schematic of RMLF reaction-front propagation where mixing is initiated by a laser and continues to travel along the foil until the layers are consumed and converted into reacted intermetallic.

[1]Materials Science and Technology Division, Lawrence Livermore National Laboratory (LLNL), Livermore, CA 94550, USA. [2]Department of Chemical Engineering and Materials Science, University of California, Davis, CA 95616, USA.

*To whom correspondence should be addressed. E-mail: kim46@llnl.gov

structures are known to arise from the interplay of thermal gradients and atomic interdiffusion, the same physical principles that govern the RMLF front propagation.

Analysis of 15-ns-resolution diffraction data (Fig. 3) reveals the initial formation of the intermetallic ordered B2 phase, NiAl, from the separate face-centered cubic (fcc) Al and $Ni_{0.91}V_{0.09}$ layers. The NiAl formed in under 300 ns after the arrival of the front. This implies that much of the interdiffusion is confined to the close proximity of the reaction front, as indicated from the time-resolved images showing the sharp onset of the cellular microstructure. The electron pulse has a 15-ns FWHM time resolution and passes through a condenser aperture so that it interacts with a round sample area that is 11.2 μm in diameter (making the obtained phase information accurate to this dimension), which is far smaller than the width of the metastable cellular formation region imaged in Fig. 2. Diffraction patterns from the unreacted bilayers appear as uniform nanocrystalline ring patterns, consistent with the ~10-nm grain diameters in the as-grown layers. Diffraction data taken only 2.3 μm ahead of the reaction front shows little indication of a phase transition. As the hot reaction front passes the sample area being probed by electrons, background intensity increases and the diffraction ring patterns become diffuse because of thermal effects of the exothermic mixing reaction. These rings are rotationally averaged to reduce noise. The film remains primarily crystalline throughout the dynamic mixing event, and this is confirmed by the uninterrupted presence of distinct rings. The timing of the hot reaction-front velocity is measured by DTEM imaging of the front of the cellular structure. When this information is combined with the diffraction data, it is evident that the solid NiAl phase begins to form immediately after the arrival of the hot reaction front.

In the reaction-front region, the intermetallic NiAl nucleates and continues to grow near the hot reaction front as the material enters a two-phase field of NiAl + liquid, consistent with known thermodynamics of Ni/Al binary mixtures (*22*). The Al atoms gain mobility because pure Al melts at the relatively low temperature of 933 K, far below the expected temperature of the reaction front. Basic calculations of heat emitted by the mixing of Al and Ni indicate that the temperature reaches ~1700 K [$\Delta H \approx 77$ kJ mol^{-1}, $c_p^{Ni/Al} = 24$ J mol^{-1} K^{-1} (*23*, *24*), where ΔH is the change in enthalpy and c_p is heat capacity at constant volume], consistent with the measured ~2% thermal expansion from the NiAl diffraction, using $\alpha = 15 \times 10^{-6}$ K^{-1} (where α is the coefficient of linear thermal expansion) (*25*). The as-grown layers are stoichiometrically close to NiAl at a 2:3 Ni-rich ratio, and the final reacted phase is NiAl (the B2 phase revealed by diffraction as in Fig. 3). Therefore, knowing that crystalline solid is present at all detectable times, it is logical that the propagating zone is moving along a thermodynamic trajectory through the NiAl + liquid phase field. This result would imply that a fraction of material at the reaction front is liquid and near equilibrium, enhancing the material transport.

The observed contrast of the cellular features may be enhanced by the liquid phase because of increased thermal diffuse scattering. On occasion, near a defect that acts as a heat sink, a small vicinity of the fully reacted foil retains cellular-type structures. The observation of such remnant features with mass-thickness contrast in high-angle annular dark field scanning TEM (fig. S5) reveals that after cooling to room temperature, the regions of narrow, dark intensity in Fig. 2 correspond to thickness modulations in the projected z direction. We confirmed the material thickness modulation using electron energy-loss spectroscopy to measure the ratio of inelastically scattered–to–elastically scattered electrons at varied positions along the specimen. The rows of increased thickness further support the idea that liquid, which may coalesce to reduce surface energy and cause this thickness modulation, is one of the phases present. This hypothesis is consistent with the thermodynamics and explains how the thickness modulations ~600-nm apart could form on

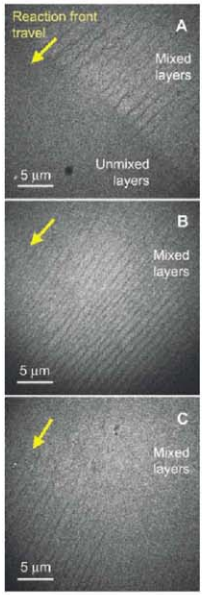


Fig. 2. RMLF mixing front is defined in (**A**), and after longer times (2 μs), the structure moves further (**B**) and is no longer present (5 μs) after the Al/Ni layers have completed mixing (**C**). The images shown here are plan-view bright field images. Dynamic snapshot images of the mixing reaction-front zone reveal a transient cellular structure.

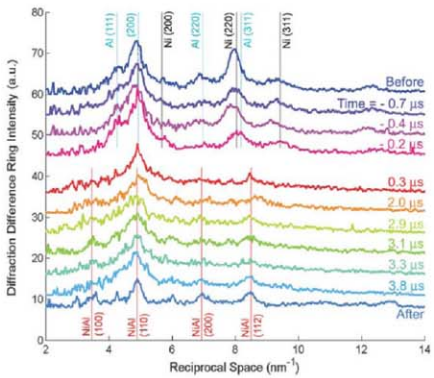


Fig. 3. Dynamic single-shot diffraction with 15-ns time resolution of regions before, during, and after the exothermic mixing reaction front has passed. The times indicated at right are in relation to the reaction front, set at $t = 0$ s. The crystal structure clearly changes from separate fcc Al/Ni and Ni/V layers to an intermetallic B2 structure NiAl phase within 300 ns after the arrival of the hot reaction front. a.u., arbitrary units.

the observed time and length scales (~100 ns judging from the micrometer-scale sharpness of the reaction front in Fig. 2A and the 13-m/s propagation speed). These time and length scales are not consistent with an alternative hypothesis that the structures arise from purely solid-state diffusive processes, which are orders of magnitude too slow, even at 1700 K (*26*).

The dark intensity between the cells fades away in the last micrograph of the series in Fig. 2, at a point in time where the reaction is long since complete. The solid solubility range of the NiAl B2 phase increases as the temperature drops (*22*), so that at room temperature nearly all of the excess Ni could be reabsorbed into a stable homogeneous B2 structure with a Ni:Al ratio close to 3:2. Post-mortem TEM examination indicates that this happens in nearly all cases, with frozen-in structures (such as those in fig. S5) being rare exceptions associated with defects. The cellular morphology is not normally present at completion.

We have obtained single-pulse nanosecond-scale TEM data in both diffraction and imaging modes, which are necessary to study the propagation and behavior of energetic nanolaminates in situ. With the use of 15-ns imaging, we have observed transient structures produced by the self-propagating high-temperature synthesis, revealing lines of mass-thickness contrast due to cellular phase formation of an ordered B2 NiAl phase and liquid. At such high formation temperatures (~1700 K), these materials are now known to exhibit transverse self-ordering reminiscent of cellular binary solidification mechanisms. We have established that the DTEM is proficient for nanosecond science in a TEM for direct nanoscale characterization of irreversible, dynamic phenomena. It is notable and exciting to find spontaneous, rapid formation of ordered structures in materials far from equilibrium, which is also an important step for essential comprehension of RMLF performance in applications.

References and Notes

1. P. E. Batson, N. Dellby, O. L. Krivanek, *Nature* **418**, 617 (2002).
2. O. L. Krivanek, N. Dellby, A. R. Lupini, *Ultramicroscopy* **78**, 1 (1999).
3. M. Haider *et al.*, *Ultramicroscopy* **75**, 53 (1998).
4. A. M. Minor, E. T. Lilleodden, E. A. Stach, J. W. Morris, *J. Electron. Mater.* **31**, 958 (2002).
5. R. Sharma, Z. Iqbal, *Appl. Phys. Lett.* **84**, 990 (2004).
6. E. D. Boyes, P. L. Gai, *Ultramicroscopy* **67**, 219 (1997).
7. M. R. Armstrong, B. W. Reed, B. R. Torralva, N. D. Browning, *Appl. Phys. Lett.* **90**, 114101 (2007).
8. P. B. Hirsch, R. W. Horne, M. J. Whelan, *Philos. Mag.* **1**, 677 (1956).
9. S. Iijima, *Nature* **354**, 56 (1991).
10. W. E. King *et al.*, *J. Appl. Phys.* **97**, 111101 (2005).
11. T. LaGrange *et al.*, *Appl. Phys. Lett.* **89**, 044105 (2006).
12. O. Bostanjoglo, R. Elschner, Z. Mao, T. Nink, M. Weingartner, *Ultramicroscopy* **81**, 141 (2000).
13. T. LaGrange, G. H. Campbell, P. E. A. Turchi, W. E. King, *Acta Mater.* **55**, 5211 (2007).
14. U. Anselmitamburini, Z. A. Munir, *J. Appl. Phys.* **66**, 5039 (1989).
15. A. J. Gavens, D. Van Heerden, A. B. Mann, M. E. Reiss, T. P. Weihs, *J. Appl. Phys.* **87**, 1255 (2000).
16. J. C. Trenkle, J. Wang, T. P. Weihs, T. C. Hufnagel, *Appl. Phys. Lett.* **87**, 153108 (2005).
17. J. Wang *et al.*, *J. Appl. Phys.* **95**, 248 (2004).
18. S. J. Zhao, T. C. Germann, A. Strachan, *J. Chem. Phys.* **125**, 164707 (2006).
19. T. P. Weihs, in *Handbook of Thin Film Process Technology*, D. A. Glocker, S. I. Shan, Eds. (Institute of Physics Publishing, Bristol, UK, 1998), p. F7.1.
20. J. S. Langer, *Rev. Mod. Phys.* **52**, 1 (1980).
21. B. E. Sundquist, *Acta Metall.* **16**, 1413 (1968).
22. M. F. Singleton, J. L. Murray, P. Nash, in vol. 1 of *Binary Alloy Phase Diagrams*, T. B. Massalski, Ed. (ASM International, Metals Park, OH, 1990), pp. 181–184.
23. C. Michaelsen, K. Barmak, T. P. Weihs, *J. Phys. D Appl. Phys.* **30**, 3167 (1997).
24. L. Perring, J. J. Kuntz, F. Bussy, J. C. Gachon, *Intermetallics* **7**, 1235 (1999).
25. Lattice-expansion measurement was determined by Reitveld analysis of the diffraction data.
26. H. Wei, X. Sun, Q. Zheng, H. Guan, Z. Hu, *Acta Mater.* **52**, 2645 (2004).

27. This work was performed under the auspices of the U.S. Department of Energy (DOE) by LLNL and supported by the Office of Science, Office of Basic Energy Sciences, Division of Materials Sciences and Engineering of the DOE under contract DE-AC52-07NA27344. J.S.K. is supported by the Lawrence Scholar Program at LLNL. The authors thank K. J. M. Blobaum and T. W. Barbee for valuable discussions, P. Ramsey for materials fabrication, and R. M. Shuttlesworth and B. J. Pyke for their support and expertise in mechanical and laser technologies.

Supporting Online Material

www.sciencemag.org/cgi/content/full/321/5895/1472/DC1
Figs. S1 to S5
Movie S1

9 June 2008; accepted 23 July 2008
10.1126/science.1161517

The Magnetic Memory of Titan's Ionized Atmosphere

C. Bertucci,[1]*† N. Achilleos,[2,3] M. K. Dougherty,[1] R. Modolo,[4] A. J. Coates,[5,3] K. Szego,[6] A. Masters,[1] Y. Ma,[7] F. M. Neubauer,[8] P. Garnier,[9] J.-E. Wahlund,[9] D. T. Young[10]

After 3 years and 31 close flybys of Titan by the Cassini Orbiter, Titan was finally observed in the shocked solar wind, outside of Saturn's magnetosphere. These observations revealed that Titan's flow-induced magnetosphere was populated by "fossil" fields originating from Saturn, to which the satellite was exposed before its excursion through the magnetopause. In addition, strong magnetic shear observed at the edge of Titan's induced magnetosphere suggests that reconnection may have been involved in the replacement of the fossil fields by the interplanetary magnetic field.

The absence of a substantial intrinsic magnetic field at Titan results in a direct interaction between the moon's chemically complex atmosphere and its plasma environment (*1*). This interaction consists of electromagnetic coupling between charged particles resulting from the ionization of Titan's atmosphere and neutral corona, and the external wind of magnetized plasma. As external plasma approaches Titan, it is mass-loaded by these locally produced ions (*2*) and its speed progressively decreases. Because the plasma is virtually collisionless, a magnetic field is frozen into the plasma, causing external field lines to drape around the moon. Magnetic field lines then pile up near the sub-flow point and stretch along the flow direction in the flanks and the downstream sector, creating an induced magnetosphere and magnetotail. In a Titan-centered frame of reference, the magnetotail consists of "away" and "toward" lobes containing, respectively, draped field lines parallel and antiparallel to the external flow (*3*, *4*). This interaction leads to the removal of ionized atmospheric constituents as they acquire the linear momentum lost by the external flow (*5*).

Titan orbits Saturn at an average distance of 20.2 Saturn radii [1 Saturn radius (R_S) = 60,268 km], spending most of its time in Saturn's partially corotating magnetospheric flow (Fig. 1A). This flow, consisting of charged particles from sources such as the E ring and Enceladus (*6*), transports Saturn's magnetic field, which encounters Titan at speeds of ~100 km/s (*1*). Hence, the angle between the magnetospheric flow and the solar photons responsible for the ionization of Titan's atmosphere—and therefore the interaction—depends on the moon's local time with respect to Saturn [Saturn local time (SLT)] (*7*). At SLT near noon, Titan can also interact with the shocked solar wind and the interplanetary magnetic field (IMF) during periods of high solar wind dynamic pressure (P_{SW}) (*8*), as the pressure balance at the magnetopause occurs at lower kronocentric distances, leaving Titan in the magnetosheath (Fig. 1B).

[1]Space and Atmospheric Physics Group, Imperial College London, The Blackett Laboratory, Prince Consort Road, London SW7 2AZ, UK. [2]Atmospheric Physics Laboratory, Department of Physics and Astronomy, University College London, Gower Street, London WC1E 6BT, UK. [3]Centre for Planetary Sciences, University College London, UK. [4]Department of Physics and Astronomy, University of Iowa, 613 Van Allen Hall, Iowa City, IA 52242–1479, USA. [5]Mullard Space Science Laboratory, University College London, Holmbury St. Mary, Dorking, Surrey RH5 6NT, UK. [6]KFKI Research Institute for Particle and Nuclear Physics, Konkoly Thege Street 29-33, Building III, H-1121 Budapest, Hungary. [7]Institute of Geophysics and Planetary Physics, University of California Los Angeles, Los Angeles, CA 90025, USA. [8]Institute for Geophysics and Meterology, Köln University, 50923 Köln, Germany. [9]Swedish Institute of Space Physics, Box 537, SE-751 21 Uppsala, Sweden. [10]Southwest Research Institute, Post Office Box 28510, San Antonio, TX 78227–1301, USA.

*Present address: Instituto de Astronomica y Fisica del Espacio, Ciudad Universitaria, Buenos Aires, Argentina.
†To whom correspondence should be addressed. E-mail: cbertucci@iafe.uba.ar

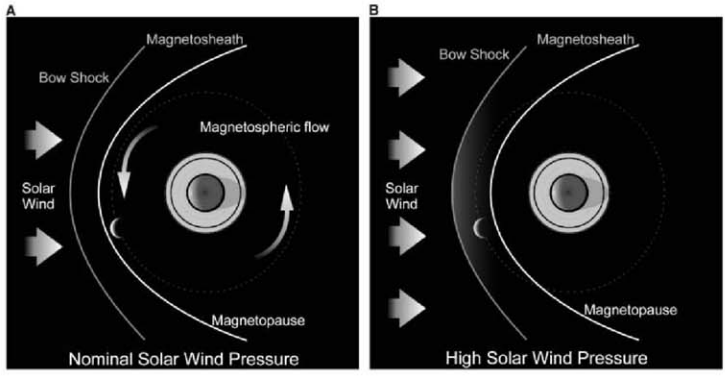


Fig. 1. Sketch describing how Titan's plasma interaction depends on solar wind pressure. Under nominal solar wind conditions, Titan interacts with Saturn's rotating magnetosphere (**A**). When solar wind pressure is high, Titan exits into Saturn's magnetosheath (**B**). Both panels indicate Titan's position during T32.

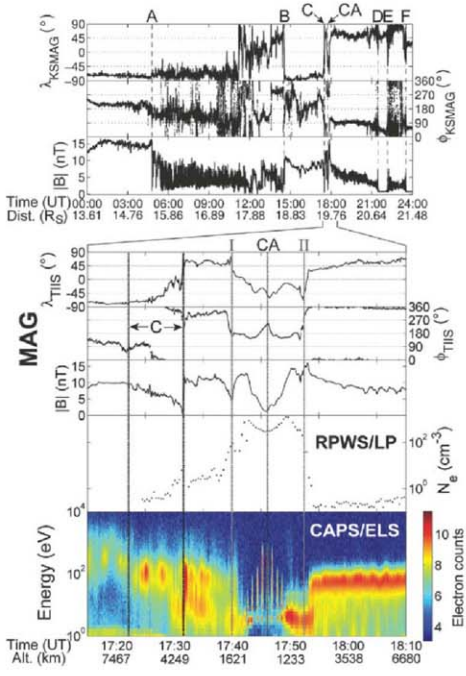


Fig. 2. (**Top**) Cassini magnetic field data on 13 June 2007 in spherical KSMAG coordinates. The encounters with the oscillating kronian magnetopause and bow shock are indicated with the letters A to C and D to F, respectively. |*B*| is the magnetic field strength. Cassini's kronocentric distance is indicated beneath the plots. (**Bottom**) Magnetic field data in spherical TIIS coordinates (from MAG), plasma density (N_e) (from RPWS/LP), and electron count rate per energy channel (from CAPS/ELS) during the T32 flyby. Magnetopause crossing C, CA, and the entry (I) and exit (II) of the fossil field region are indicated. Cassini's altitude above Titan is indicated beneath the plots.

During every in situ observation made from Voyager 1 (*3*) through Cassini close flyby 31 (T31), Titan was inside the kronian magnetosphere. However, during T32 on 13 June (day 164) 2007, Cassini encountered Titan while it was in the shocked solar wind. Observations of the plasma environment showed layers of remnant kronian magnetic fields to which Titan had been exposed a few minutes before, revealing how this field interacts with the IMF.

T32 took place at 13.6 hours SLT, close to the Saturn-Sun line, when the spacecraft was outbound from the kronian system. In the Titan-centered frame, Cassini's trajectory was almost parallel to the Saturn-Sun direction and north of the moon's orbital plane. Closest approach (CA) occurred at 17:46:32 universal time (UT) (*9*) at an altitude of 975 km over the north pole. As a result, Cassini first explored the near-Saturn side of Titan's induced magnetosphere, then flew through the collisional ionosphere and emerged on the side facing away from the planet.

The T32 encounter occurred after a series of compressions and expansions of Saturn's magnetopause (at speeds much higher than Cassini's) in response to strong variations in P_{SW}. These oscillations are noticeable in the Cassini magnetometer (MAG) (*10*) data in the spherical KSMAG coordinate system (*11*) aligned with Saturn's magnetic dipole (Fig. 2, top panel). Early on day 164, Cassini was immersed in a typical north/south kronian magnetic field. Then the magnetosphere contracted because of an increase in P_{SW}. The receding magnetopause passed Cassini around 04:47 (point A) at a distance of ~15.4 R_S from Saturn. A model of the magnetopause based on pressure balance (*12*, *13*) suggests that P_{SW} was ~0.08 nPa (more than five times the average value). From that moment, Titan was also outside the magnetosphere. Furthermore, if the same dependence on P_{SW} is applied to the bow shock (*14*), Titan was in the supersonic solar wind at the time of A. After A, Cassini spent almost 10 hours outside Saturn's magnetosphere. Meanwhile, the magnetopause stopped contracting and expanded once again, reaching Cassini around 14:32 (point B). Cassini was then within the magnetosphere for ~3 hours until it encountered, one last time, a retreating magnetopause around 17:25 (point C), more than 20 min before CA. The magnetic field disturbance generated by Titan was observed by Cassini ~10 min after C. After ~11:00, the IMF was mainly northward, leading to strong magnetic shear at crossings B and C.

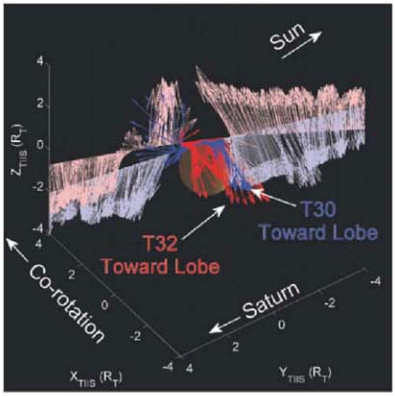


Fig. 3. Magnetic field measurements along the trajectories of flybys T30 (blue) and T32 (red) in TIIS coordinates. The draped magnetic fields of Titan's induced magnetosphere are highlighted in bright colors. Light-colored blue and red arrows indicate the unperturbed ambient field.

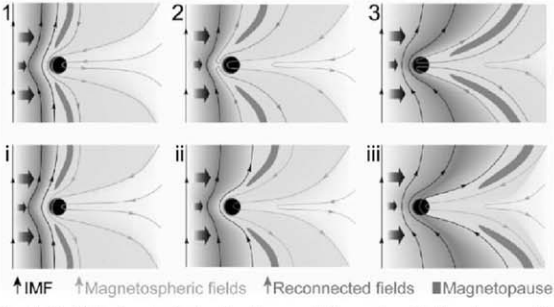


Fig. 4. Simplified schematics showing, from the same initial scenario and in three stages, possible reconfigurations of the magnetic field near Titan during T32. **(Top)** Saturn's magnetopause opens as a result of reconnection between the kronian fossil fields and the IMF (1), and reconnected field lines are carried downstream (2 and 3). **(Bottom)** The absence of reconnection (i) results in a closed magnetopause that sweeps across Titan's induced magnetosphere (ii and iii). The shocked solar wind travels from left to right. The formation of the fossil fields and possible tail reconnection are not shown because of the geometrical complexity of the process.

The combined capabilities of Cassini MAG, the Radio and Plasma Wave Science instrument (RPWS) (*15*), and the Cassini Plasma Spectrometer/Electron Spectrometer (CAPS/ELS) (*16*) provide a detailed description of the plasma near Titan (Fig. 2, bottom panel). Magnetic field data are shown in the spherical Titan interaction (TIIS) coordinate system defined from the nominal kronian corotation flow (*4*, *17*). In Cassini's frame of reference, the magnetopause crossing C displayed a thick boundary layer, where the magnetospheric and the magnetosheath plasmas coexisted, whereas the magnetic field rotated northward by 156°. These signatures are typical of local magnetopause reconnection events previously observed at Saturn (*18*). Also within C, ELS and the RPWS Langmuir probe (LP) detected (from 17:30) a less energetic (10-eV) electron population that probably originated from Titan. After C (between 17:32 and 17:38, and after 17:53), the magnetic field was predominantly northward, confirming that Titan was in the IMF. The mixed magnetosphere/magnetosheath plasma signature continued until 17:38.

Around 17:41 and 17:53 (altitudes of 1400 and 1740 km, respectively), the sudden drop in 100- to 1000-eV electron count rates indicates that the external flow was strongly decelerated and deflected near Titan, as its cold plasma began to dominate. Simultaneously, the plasma density increased above 100 cm^{-3} and the frequency of collisions became comparable to the ion gyrofrequency. As a result, only electrons remained magnetized and deposited the magnetic field in the induced magnetosphere via convective pileup. These electrons had lost most of their momentum, making the magnetic flux tube convection time extremely long as compared to that at higher altitudes.

From 17:43:30 to 17:49:00 (below an altitude of 1100 km), extremely weak fields indicate that Cassini entered Titan's collisional ionosphere, where magnetic diffusion dominated over convec-

tion. A minimum value of 1.25 nT was measured a few seconds before CA and below the ionospheric peak around 17:42 and 17:50 (altitudes of 1270 and 1200 km, respectively).

Between approximately 1200 and 1600 km altitude (points I and II), the draped magnetic field was antiparallel to the corotation direction (the toward lobe). Because Cassini was at northern latitudes, this lobe corresponds to the draping of a field that was initially southward. However, the IMF surrounding Titan was clearly northward, and there is no possibility that such an orientation could lead to a toward lobe at northern latitudes. A comparison between the average field in Titan's magnetotail during T32 with that in the toward lobe during three north latitude flybys with almost identical geometries (T28, T29, and T30) inside Saturn's magnetosphere shows that these directions differed by as little as 28°. The fields in the toward lobes of flybys T32 and T30 (Fig. 3) display high similarity (8° between averages).

In the absence of an internal field, the fields just above Titan's collisional ionosphere are draped kronian magnetospheric fields deposited within Titan's induced magnetosphere before the moon's magnetosheath excursion. These fields are effectively frozen or "fossilized" in the near-Titan noncollisional plasma as a result of mass loading by cold ions [notably N_2^+ (*19*)] from Titan's exosphere below 1600 km, dramatically reducing the transit speed of magnetic flux tubes. As Titan entered the magnetosheath, it encountered shocked solar wind traveling along the Sun-Saturn direction. This flow transported IMF field lines, which piled up and draped around the layers of fossilized fields, as shown by the monotonic change in the magnetic field strength after 17:53. Thus, draped field lines at higher altitudes moved at higher speeds, reflecting more recently sampled magnetic environments.

The detection of kronian fossil fields during T32 occurred because their convection time was longer than the time that Titan was exposed to the IMF. In the early hours of day 164, Titan was immersed in the magnetosheath longer than Cassini (9 hours, 45 min), because the spacecraft was inside Titan's orbit. For the same reason, Titan was within Saturn's magnetosphere for a shorter time than the interval from B to C. However, this was long enough to replenish Titan's induced magnetosphere with kronian fields after the first magnetosheath excursion. Hence, if Titan was within Saturn's magnetosphere for some time between B and C, and if the replenishment time was the same for the IMF and magnetospheric fields, the time spent by Titan in Saturn's magnetosphere between B and C should have been longer than that between C and CA. As a result, the lifetime of the kronian fossil fields was between ~20 min and ~3 hours. These conclusions support previous theoretical estimations within Saturn's magnetosphere (*20*).

The fossil fields could have been removed via diffusion into the collisional ionosphere, convection around the ionosphere into the magnetotail, and reconnection with the IMF. The magnetic shear at locations I and II (111° and 162°, respectively) and the magnetic field variance suggest that reconnection could have occurred during T32. In this scenario (Fig. 4, top panels), Titan could have opened Saturn's magnetopause, and the field reconfiguration could have been similar to that proposed for disconnection events at comets (*21*). Initial real-time modeling supports this interpretation.

However, previous simulations (*20*) indicate that at these altitudes, the plasma density was too low for electrons to be demagnetized. In such a case (Fig. 4, bottom panels), reconnection would not have occurred, and the fossilized fields could have been either diffused into the collisional ionosphere or transported downstream by ambipolar electric fields (*22*) without affecting the magnetic structure of Saturn's magnetopause.

The bow-shock crossings D, E, and F at kronocentric distances of 20.8, 21.0, and 21.4 R_S (Fig. 2, top panel) suggest magnetopause standoff distances around 16 R_S. Hence, it is likely that after T32, Titan remained within Saturn's magnetosheath at least until 23:30, when the fossil fields should have been entirely removed.

References and Notes

1. F. M. Neubauer, D. A. Gurnett, J. D. Scudder, R. E. Hartle, in *Saturn*, T. Gehrels, M. S. Matthews, Eds. (Univ. of Arizona Press, Tucson, AZ, 1984), pp. 760–787.
2. J.-E. Wahlund *et al.*, *Science* **308**, 986 (2005).
3. N. F. Ness, M. H. Acuña, K. W. Behannon, *J. Geophys. Res.* **87**, 1369 (1982).
4. H. Backes *et al.*, *Science* **308**, 992 (2005).
5. K. Szego *et al.*, *Geophys. Res. Lett.* **34**, L24503 (2007).
6. M. G. Kivelson, *Science* **311**, 1391 (2006).
7. SLT is measured from the noon meridian in the direction of Saturn's rotation.
8. $P_{SW} = \rho v^2$, where ρ and v are the solar wind density and speed, respectively.
9. All times are expressed in UT.
10. M. K. Dougherty *et al.*, *Space Sci. Rev.* **114**, 331 (2004).
11. In the Saturn-centered spherical solar magnetic coordinate system (KSMAG), the polar axis coincides with planet's magnetic axis and λ_{KSMAG} and ϕ_{KSMAG} are, respectively, the magnetic field latitude and azimuth ($\phi_{KSMAG} = 0$ corresponds to the magnetic noon meridian).
12. $R_{MP} = P_{SW}^{-1/4.3}$, where R_{MP} is the magnetopause's standoff distance.
13. C. S. Arridge *et al.*, *J. Geophys. Res.* **111**, A11227 (2006).
14. A. Masters *et al.*, *J. Geophys. Res.* 10.1029/2008JA013276 (2008).
15. D. A. Gurnett *et al.*, *Space Sci. Rev.* **114**, 395 (2004).
16. D. T. Young *et al.*, *Space Sci. Rev.* **114**, 1 (2004).
17. In the Titan-centered spherical ionospheric interaction coordinate system (TIIS), the equatorial plane contains the direction to Saturn (Y_{TIIS}) and that of the nominal corotation flow (X_{TIIS}). λ_{TIIS} and ϕ_{TIIS} are, respectively, the latitude and azimuth (measured from X_{TIIS} toward Y_{TIIS}).
18. H. J. McAndrews *et al.*, *J. Geophys. Res.* **113**, A04210 (2008).
19. R. E. Hartle *et al.*, *Geophys. Res. Lett.* **33**, L08201 (2006).
20. H. Backes, thesis, University of Cologne, Germany (2005).
21. M. B. Niedner, J. C. Brandt, *Astrophys. J.* **223**, 655 (1978).
22. A. J. Coates *et al.*, *Geophys. Res. Lett.* **34**, L24505 (2007).

23. We thank A. Balogh, T. Horbury, W. S. Kurth, C. Mazelle, S. J. Schwartz, and C. T. Russell for useful discussions, and N. Powell for artwork. C.B. was supported by a Science and Technology Facilities Council Postdoctoral Fellowship.

29 April 2008; accepted 28 July 2008
10.1126/science.1159780

Postseismic Relaxation Along the San Andreas Fault at Parkfield from Continuous Seismological Observations

F. Brenguier,[1,2]* M. Campillo,[2] C. Hadziioannou,[2] N. M. Shapiro,[1] R. M. Nadeau,[3] E. Larose[2]

Seismic velocity changes and nonvolcanic tremor activity in the Parkfield area in California reveal that large earthquakes induce long-term perturbations of crustal properties in the San Andreas fault zone. The 2003 San Simeon and 2004 Parkfield earthquakes both reduced seismic velocities that were measured from correlations of the ambient seismic noise and induced an increased nonvolcanic tremor activity along the San Andreas fault. After the Parkfield earthquake, velocity reduction and nonvolcanic tremor activity remained elevated for more than 3 years and decayed over time, similarly to afterslip derived from GPS (Global Positioning System) measurements. These observations suggest that the seismic velocity changes are related to co-seismic damage in the shallow layers and to deep co-seismic stress change and postseismic stress relaxation within the San Andreas fault zone.

Information about the stress variations in deeper parts of continental faults can be obtained by studying source properties of microearthquakes (*1*). Changes in seismic velocities measured by using repeated natural and active seismic sources can also provide information about rock damage and healing at depth after large earthquakes (*2*, *3*) or about stress changes in seismogenic zones (*4*). The main limitation of these types of measurements, however, is the episodic nature of their seismic sources, which prevents continuous monitoring of crustal properties.

We used continuous measurements of ambient seismic noise to recover continuous variations of seismic velocities within the crust along the San Andreas fault (SAF) near Parkfield, California. With this approach, the cross-correlation function of ambient seismic noise computed between a pair of receivers converges toward the response of Earth between the receivers (the so-called Green's function). Essentially this function represents the seismogram that would be recorded at one of the receivers if a source were acting at the second (*5*, *6*). The

temporal evolution of the crust is then tracked by computing cross-correlation functions at different dates for the same receiver pair and measuring the changes between the correlation functions (*7–9*).

To monitor variations in seismic velocity along the SAF at Parkfield, we used more than 5 years of continuous seismic noise data recorded by 13 short-period seismological stations of the Berkeley High Resolution Seismic Network (HRSN) (*10*). These stations are installed in boreholes at depths of 60 to 300 m, thus reducing locally generated noise and effects of temperature variations and precipitation (Fig. 1). We analyzed data from January 2002 to October 2007, spanning the times of two major earthquakes that occurred within a 100-km radius of Parkfield: the moment magnitude (M_w) = 6.5 San Simeon earthquake of 22 December 2003, whose epicenter was located 60 km west of Parkfield, and the M_w = 6.0 Parkfield earthquake of 28 September 2004. For every possible pair combination of stations, we computed the daily cross-correlation of seismic noise by using the procedure of (*11*), yielding 91 × 2140 days = 194,740 cross-correlation and auto-correlation time functions. A reference Green function (RGF) was computed for each station pair by stacking the daily cross-correlations for the entire 2140-day period (*12*). The velocity changes were then determined by measuring time delays between the RGF and 30-day stacks of cross-correlation functions in the frequency range from 0.1 to 0.9 Hz (*9*, *12*, *13*) (Fig. 2B). If the medium experiences a spatially homogeneous relative seismic velocity change, $\Delta v/v$, the relative travel time shift ($\Delta t/t$) between a perturbed and reference Green function is independent of the lapse time (t) at which it is measured and $\Delta v/v = -\Delta t/t$ = constant. Therefore, when computing a local time shift, Δt, between the reference and a chosen cross-correlation function in a short window centered at time t, we would expect that Δt should be a linear function of t. By measuring the slope of the travel time shifts Δt as a function of time t, we then estimated the relative time perturbation ($\Delta t/t$), which is the opposite value of the medium's relative velocity change ($\Delta v/v$). The 30-day stacked correlations shown in Fig. 2A exhibit variations because of the seasonal pattern of the location of noise sources (*14*, *15*). Because these seasonal variations mainly affect the direct waves, we did not make differential time measurements for these waves. We also investigated the accuracy of the station clocks by analyzing the temporal symmetry of the correlation functions (*16*) and correcting for the detected errors (*12*). Lastly, following (*9*), we averaged the measured time shifts for each time t over all station pairs to increase the measurement accuracy.

After the San Simeon earthquake, the seismic velocity along the SAF at Parkfield decreased by 0.04% (Fig. 3). This is consistent with measurements using active sources and fault guided waves that are associated with other earthquakes (*2*, *3*, *17*). Creepmeter and Global Positioning System (GPS) measurements show that there was no substantial slip detected along the SAF in the Parkfield area after the San Simeon earthquake (*18*). This suggests that the velocity change we detected may be related to co-seismic damage in the shallow layers caused by strong ground shaking (~0.15 g) from this quake. By 7 months after the quake, velocities in the Parkfield area appear to have returned to their pre-earthquake levels.

Kinematic and dynamic rupture inversions as well as GPS and INSAR (Interferometric Synthetic Aperture Radar) measurements showed that the Parkfield mainshock released a maximum stress of 10 Mpa and that the average slip was about 0.5 m (*19*). The Parkfield mainshock was also followed by postseismic afterslip that is still ongoing and broadly distributed between the surface and a depth of 12 km (*20*, *21*). Immediately after the Parkfield earthquake, velocities decreased by 0.08%, and postseismic velocities remained low for almost 3 years (Fig. 3). The long-term decay of the relative velocity perturbation was very similar to the relaxation curve associated with the along-fault displacement deduced from GPS measurements (*21*, *22*). Therefore, our hypothesis is that the evolution of the observed seismic velocity changes after the Parkfield earthquake was governed by the postseismic stress relaxation within deeper parts of the fault zone and the surrounding region.

Observation of nonvolcanic tremors in the vicinity of the Parkfield area supports this hypothesis (Fig. 3). We considered the 30-day averaged rate of tremor activity in the Cholame-Parkfield region computed by using continuous records from the HRSN for the period 2002 through 2007. These tremors are estimated to have occurred between 20- and 40-km depths (*23*), similarly to the episodic tremor and slip phenomena on subduction zones (*24*, *25*). There is clear evidence of triggering of tremor activity by both San-Simeon and Parkfield earthquakes. After the Parkfield earthquake, tremor activity remained elevated and has yet to return to its pre-event level similarly to the seismic velocity changes. This observation supports our hypothesis that both seismic velocity changes and tremor activity after the Parkfield earthquake are related to postseismic stress relaxation and corresponding slow slip. We also propose that the increased nonvolcanic tremor activity after the San Simeon earthquake may be related to slow slip at depth in response to small stress variations induced by the passing of seismic waves from the M_w = 6.5 event (*26*).

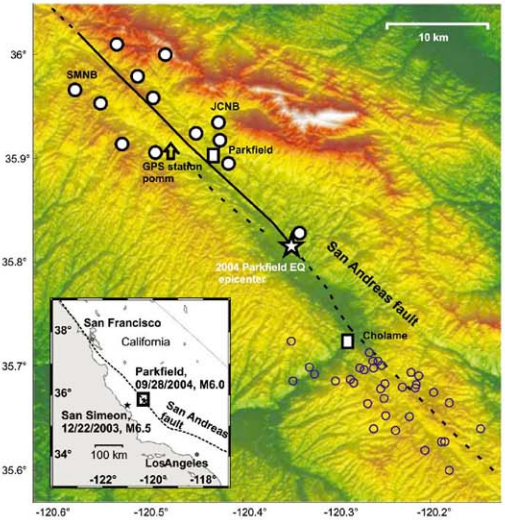


Fig. 1. Location of the HRSN (white and black circles) near Parkfield, California, and location of the 2003 San Simeon and 2004 Parkfield earthquakes. The black solid line indicates the surface projection of the 2004 Parkfield earthquake rupture and afterslip extent. The blue circles indicate the epicenters of nonvolcanic tremors detected by (*23*). The black box on the inset image corresponds to the studied area. The DEM plot was obtained from (*27*, *28*). EQ indicates earthquake.

[1]Sismologie, Institut de Physique du Globe de Paris (IPGP) and CNRS, 4 Place Jussieu, 15252 Paris, France. [2]Laboratoire de Géophysique Interne et Tectonophysique, Université Joseph Fourier and CNRS, 38041 Grenoble, France. [3]Berkeley Seismological Laboratory, University of California Berkeley, 215 McCone Hall, Berkeley, CA 94760, USA.

*To whom correspondence should be addressed. E-mail: fbrenguI@ipgp.jussieu.fr

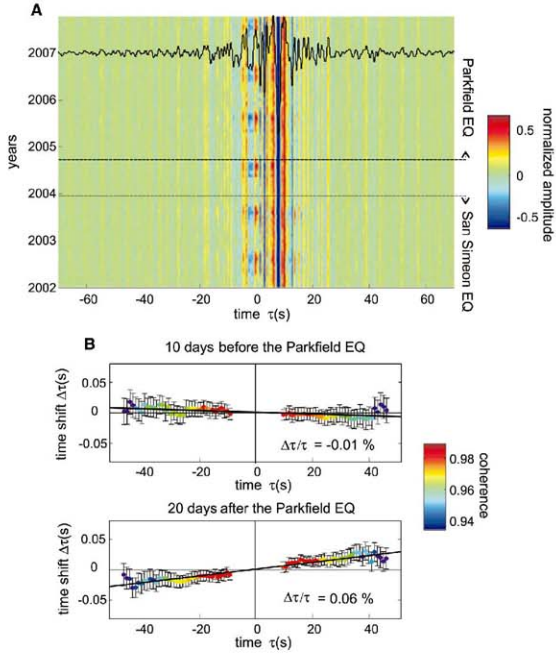


Fig. 2. Relative travel-time change measurements ($\Delta\tau/\tau$). **(A)** Thirty-day stacked cross-correlation functions (CCF) for receiver pair JCNB-SMNB. The black curve represents the reference stacked cross-correlation function. The CCFs are filtered between 0.1 and 0.9 Hz and normalized in amplitude. **(B)** Time shifts averaged over 91 receiver pairs and coherence measured between the reference stacked and 30-day stacked cross-correlation functions (frequency band, 0.1 to 0.9 Hz).

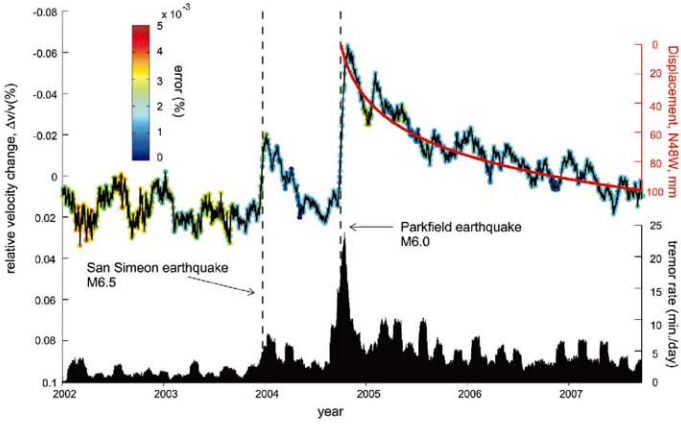


Fig. 3. Seismic velocity changes, surface displacements from GPS, and tremor activity near Parkfield. The red curve represents the postseismic fault-parallel displacements along the San Andreas fault as measured by GPS at station pomm (Fig. 1) (*29*). The tremor rates are averaged over a centered 30-day-length moving time window.

Differences in the evolution of seismic velocities after the San Simeon and the Parkfield earthquakes indicate that two different physical mechanisms may be responsible for the changes in crustal properties: (i) damage of shallow layers and fault zone caused by the strong ground shaking and (ii) co-seismic stress change followed by the postseismic relaxation. These results demonstrate that measuring small velocity perturbations from correlations of seismic noise can be a useful tool for studying the continuous time evolution of the stress regime in the vicinity of seismogenic faults.

References and Notes

1. B. Allmann, P. Shearer, *J. Geophys. Res.* **112**, B04305 (2007).
2. J. E. Vidale, Y. Li, *Nature* **421**, 524 (2003).
3. Y. Li, J. E. Vidale, S. Day, D. Oglesby, E. Cochran, *Bull. Seismol. Soc. Am.* **93**, 854 (2003).
4. F. Niu, P. G. Silver, T. M. Daley, X. Cheng, E. L. Majer, *Nature* **454**, 204 (2008).
5. N. M. Shapiro, M. Campillo, L. Stehly, M. H. Ritzwoller, *Science* **307**, 1615 (2005).
6. M. Campillo, *Pure Appl. Geophys.* **163**, 475 (2006).
7. C. Sens-Schönfelder, U. Wegler, *Geophys. Res. Lett.* **33**, L21302 (2006).
8. U. Wegler, C. Sens-Schönfelder, *Geophys. J. Int.* **168**, 1029 (2007).
9. F. Brenguier *et al.*, *Nat. Geosci.* **1**, 126 (2008).
10. More details concerning the Parkfield HRSN can be found at http://seismo.berkeley.edu/bdsn/hrsn.overview.html.
11. F. Brenguier, N. M. Shapiro, M. Campillo, A. Nercessian, V. Ferrazzini, *Geophys. Res. Lett.* **34**, L02305 (2007).
12. Materials and methods are available as supporting material on *Science* Online.
13. G. Poupinet, W. L. Ellsworth, J. Frechet, *J. Geophys. Res.* **89**, 5719 (1984).
14. L. Stehly, M. Campillo, N. Shapiro, *J. Geophys. Res.* **111**, B10306 (2006).
15. S. Kedar, F. H. Webb, *Science* **307**, 682 (2005).
16. L. Stehly, M. Campillo, N. Shapiro, *Geophys. J. Int.* **171**, 223 (2007).
17. J. L. Rubinstein, G. Beroza, *Geophys. Res. Lett.* **32**, L14313 (2005).
18. More details concerning the U.S. Geological Survey (USGS) deformation network at Parkfield can be found at http://earthquake.usgs.gov/research/parkfield/deform.php.
19. S. Ma, S. Custódio, R. Archuleta, P. Liu, *J. Geophys. Res.* **113**, B02301 (2008).
20. K. M. Johnson, R. Burgmann, K. Larson, *Bull. Seismol. Soc. Am.* **96**, S321 (2006).
21. A. Freed, S. Ali, R. Bürgmann, *Geophys. J. Int.* **169**, 1164 (2007).
22. I. Johanson, E. Fielding, F. Rolandone, R. Burgmann, *Bull. Seismol. Soc. Am.* **96**, S269 (2006).
23. R. M. Nadeau, D. Dolenc, *Science* **307**, 389 (2005); published online 9 December 2004 (10.1126/science.1107142).
24. K. Obara, *Science* **296**, 1679 (2002).
25. G. Rogers, H. Dragert, *Science* **300**, 1942 (2003); published online 8 May 2003 (10.1126/science.1084783).
26. J. Rubinstein *et al.*, *Nature* **448**, 579 (2007).
27. The digital elevation model (DEM) was obtained from the USGS National Map Seamless Server, http://seamless.usgs.gov.
28. J. F. Luis, *Comput. Geosci.* **33**, 31 (2007).
29. More details concerning the USGS GPS network at Parkfield can be found at http://quake.usgs.gov/research/deformation/twocolor/pkf_continuous_gps.html.

30. All of the data used in this study came from the Parkfield HRSN and were collected by the Berkeley Seismological Laboratory (BSL) with support from the USGS under grant 07HQAG0014. We are grateful to the BSL staff. We thank P. Bernard and J. P. Gratier for discussions, O. Coutant for providing us with the doublet analysis code, and G. Moguilny for maintaining the Cohersis cluster. We acknowledge the December 2007 joint BSL/IPGP workshop. This work was supported by Agence Nationale de la Recherche (France) under contracts 05-CATT-010-01 (PRECORSIS) and ANR-06-CEXC-005 (COHERSIS); by NSF under grants EAR-0537641 and EAR-0544730; and by USGS under grant 06HQGR0167. This is Institut de Physique du Globe de Paris (IPGP) contribution no. 2393.

Supporting Online Material

www.sciencemag.org/cgi/content/full/321/5895/1478/DC1
Materials and Methods
Fig. S1

27 May 2008; accepted 17 July 2008
10.1126/science.1160943

Atmospheric Warming and the Amplification of Precipitation Extremes

Richard P. Allan[1]* and Brian J. Soden[2]

Climate models suggest that extreme precipitation events will become more common in an anthropogenically warmed climate. However, observational limitations have hindered a direct evaluation of model-projected changes in extreme precipitation. We used satellite observations and model simulations to examine the response of tropical precipitation events to naturally driven changes in surface temperature and atmospheric moisture content. These observations reveal a distinct link between rainfall extremes and temperature, with heavy rain events increasing during warm periods and decreasing during cold periods. Furthermore, the observed amplification of rainfall extremes is found to be larger than that predicted by models, implying that projections of future changes in rainfall extremes in response to anthropogenic global warming may be underestimated.

Predicting and adapting to changes in the global water cycle expected to result from global warming presents one of the greatest challenges to humanity. Projections of tropical precipitation through this century anticipate increases in moist equatorial regions and indications of drying over the already-arid subtropics (*1*–*6*), changes consistent with theoretical considerations (*7*–*9*). Low-level moisture rises with temperature at about 7%/K, as expected from the Clausius Clapeyron equation (*10*, *11*), fueling comparable rises in heavy precipitation events driven by moisture convergence (*8*). Mean precipitation and evaporation are constrained by the slower rises in atmospheric radiative cooling to the surface (*7*, *12*, *13*). This leads to a decline in precipitation away from the convectively driven regimes that the models achieve through reduced water vapor mass flux and wind stress associated with a weakening of the Walker circulation (*14*). Observational evidence supports the findings that moist regions are becoming wetter and dry regions drier (*3*, *15*, *16*), but the overall response of the models to the current warming trend appears underestimated (*11*, *15*–*17*) and the cause of this discrepancy may affect the fidelity of climate predictions.

Present-day changes in the tropical water cycle are dominated by the periodic warming and moistening associated with El Niño Southern Oscillation (ENSO). Figure 1 shows that warm ENSO events (positive Nino-3 index) are associated with higher column water vapor and precipitation, whereas the reverse is true for cold events. This variability provides a means for testing hypotheses regarding how precipitation responds to a warmer climate. The contrasting mechanisms involved in driving heavy and light precipitation necessitate the examination of daily data (*5*, *18*): We compared daily precipitation from the Special Sensor Microwave Imager (SSM/I) over the tropical oceans (*11*) to multiple Coupled Model Intercomparison Project 3 (CMIP3) models (*19*) forced with present-day sea surface temperature (SST) and with projected greenhouse gas concentrations for the 21st century.

The SSM/I data resolution was degraded to 2.5° by 2.5°, a resolution that is more comparable with climate model grids. Each month of daily precipitation maps was partitioned into 12 bins ranging from the lightest 10% up to the heaviest 1% [supporting online material (SOM) text]. Because climate models struggle to simulate the observed distribution of rainfall intensities (*20*), we calculated changes in precipitation frequency in each bin separately for the satellite data and for each model. Bin boundaries were calculated from 1 year of daily data for the models and the satellite data; the year chosen does not alter the boundary rainfall intensity by more than 10% (table S1). Deseasonalized anomalies in the frequency of precipitation were calculated for each bin. The percentage changes in precipitation frequency, relative to the mean frequency for each bin, were thus calculated for each bin every month.

Figure 2 presents the percentage changes in precipitation frequency in each bin for the SSM/I data and models. These are comparable with results from the SSM/I data for 10 exact 10% bins at the original and the degraded resolutions (fig. S1) and for individual models (fig. S2). There is a coherent variability in observed very heavy precipitation, with higher frequencies associated with warm El Niño events (1988, 1991, 1997–8, and 2002–3) and lower frequencies with cold La Niña events (1989, 1996, and 1999–2000). The model ensemble mean (Fig. 2B) shows qualitative agreement with the satellite data for the heaviest rainfall bins; agreement

[1]Environmental Systems Science Centre, University of Reading, Berkshire RG6 6AL, UK. [2]Rosenstiel School of Marine and Atmospheric Sciences, University of Miami, FL 33149, USA.
*To whom correspondence should be addressed. E-mail: r.p.allan@reading.ac.uk

is improved further by sampling only the wettest 20% of all model grid boxes (fig. S3), essentially masking out much of the light rainfall in the models. However, whereas the frequency of the light rainfall events (10 to 20% bin) tends to be anti-correlated with the frequency of very heavy precipitation (95 to 99% bin) in the satellite data ($r = -0.51$), the reverse is true for the models ($r = 0.72$). Instead, a greater occurrence of very heavy rainfall events in the models is at the expense of the heavy rainfall events (70 to 80% bins) ($r = -0.90$), at odds with the variations captured by the satellite data.

Observational evidence (*11*, *17*) suggests that mean precipitation and evaporation are currently rising at the rate expected from the Clausius Clapeyron equation. Do observations and model simulations of precipitation extremes also follow this simple thermodynamic relationship? To test this hypothesis, we performed a Clausius Clapeyron (C-C) experiment whereby data from 1 year of daily precipitation fields, from the atmospheric component of the Geophysical Fluid Dynamics Laboratory (GFDL) coupled climate model version 2.1 (*21*), were perturbed by 7% times the local monthly mean Hadley Centre Sea Ice and Sea Surface Temperature (HadISST)–observed SST anomaly (*22*). This allows us to identify what component of the precipitation response in models and observations can be explained from C-C considerations alone (Fig. 2C). The very heavy rainfall response shows similarity to the satellite data and the models, consistent with theory (*8*). However, the changing frequency of precipitation in the heavy, moderate, and light bins for the C-C experiment cannot explain the observed or simulated variability.

Figure 3 shows observed and simulated time series of the percentage anomalies in precipitation frequencies for the three heaviest precipitation bins.

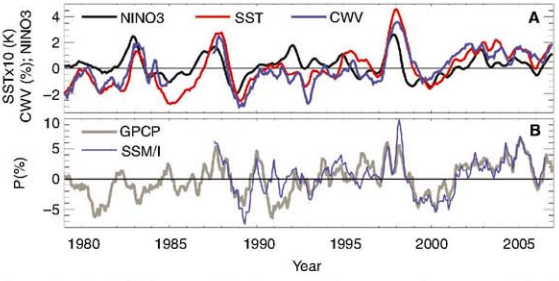


Fig. 1. Time series of **(A)** Nino-3 ENSO index (SST anomalies for 90° to 150°W, 5°S to 5°N region) and deseasonalized tropical ocean (30° to 30°N) mean anomalies of SST from HadISST (*22*) and column-integrated water vapor (CWV) from passive microwave satellite data (Scanning Multichannel Microwave Radiometer, SSM/I) and **(B)** precipitation (P) from GPCP and SSM/I (*11*).

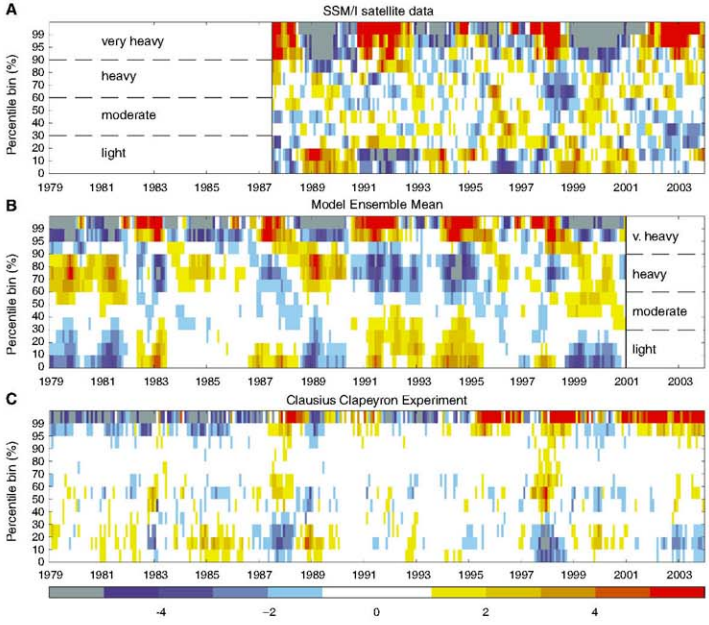


Fig. 2. Percentage anomalies of precipitation frequency in bins of rainfall intensity for **(A)** SSM/I data, **(B)** climate model ensemble mean, and **(C)** C-C experiment (precipitation increases at 7%/K).

There are substantial fluctuations in precipitation frequency anomalies in some models, but the model ensemble mean and satellite observations demonstrate consistent variance (fig. S4) and coherent variability in the heaviest rainfall bin (Fig. 3A) with the El Niño events of 1987–88, 1991, 1994–95, and 1997–98, coinciding with increased frequency of the heaviest precipitation. In the second-heaviest precipitation bin (Fig. 3B), the model ensemble response displays positive correlation with the C-C experiment ($r = 0.62$) as does the satellite data ($r = 0.61$), albeit with greater variance (fig. S4).

The response of precipitation frequency to changes in SST were quantified for the present-day variability and compared with the response to global warming. Linear fits between precipitation frequency in each bin and tropical mean SST were constructed (Fig. 4A). The heaviest precipitation bin displays a dependence on SST two to three times greater than the model. The observed relationship is robust to the processing applied to the satellite data; only by sampling the wettest 20% of all model grid boxes does the sensitivity of precipitation frequency to SST begin to resemble the satellite observations (fig. S5). Although moderate and light precipitation bins display a weak negative dependence on SST for the SSM/I data, the model response is more strongly negative for heavy rainfall (60 to 90%), and the frequency of light precipitation (below 30th percentile) increases with SST. Because the statistical significance is weak (only the 95 to 99 and 99 to 100% bins produce correlations with SST above the 95% significance level for both models and satellite data), the relationships are confirmed independently by compositing El Niño and La Niña months separately. Figure 4B shows essentially the same differences between models and observations, albeit with a stronger sensitivity.

It is possible that changes in atmospheric circulation associated with ENSO may affect the calculated relationships because during El Niño there tends to be a drying over land and moistening over oceans (*16*, *23*). To test this hypothesis, we recalculated the relationships in Fig. 4, A and B, for the entire tropics (land and ocean) by using model data. Because differences to the ocean-only calculations are small, we can conclude that enhanced oceanic ascent during El Niño plays only a minor role in determining relationships between precipitation frequency and SST.

Can the changes in precipitation frequency be explained by thermodynamic considerations? The linear regression between precipitation frequency and changes in SST for the C-C experiment (Fig. 4C) shows reduced occurrence of light precipitation and increased frequency of very heavy precipitation, as expected were precipitation to be simply scaled by a constant factor (SOM text and fig. S6). This response is consistent with the model-simulated response of the heaviest precipitation but does not capture the changes for other rainfall intensities (Fig. 4A). The SSM/I response of the very heavy precipitation frequency appears larger than expected from C-C changes; this remains theoretically possi-

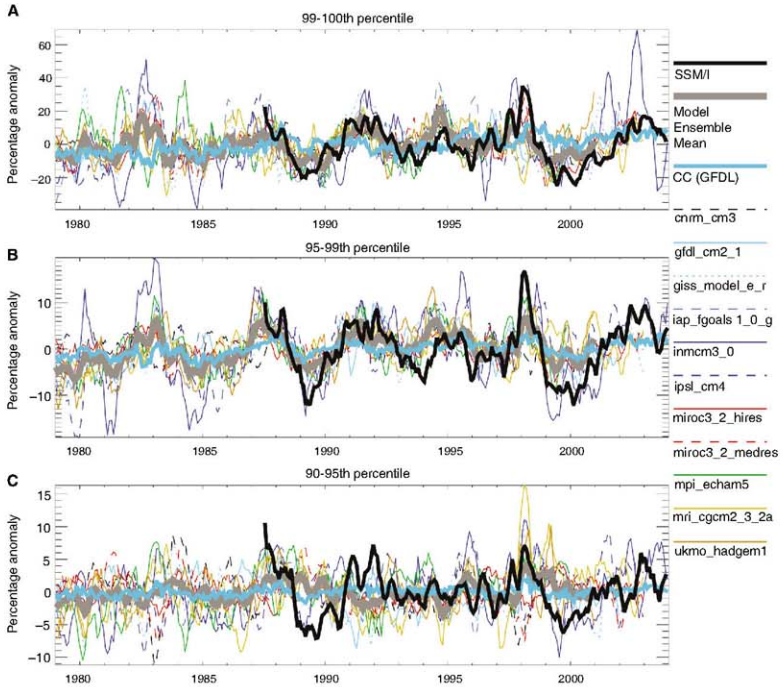


Fig. 3. Percentage anomalies in precipitation frequencies for the **(A)** 99th to 100th percentile, **(B)** 95th to 99th percentile, and **(C)** 90th to 95th percentile bins for individual models, the model ensemble mean, the C-C experiment, and the SSM/I satellite data.

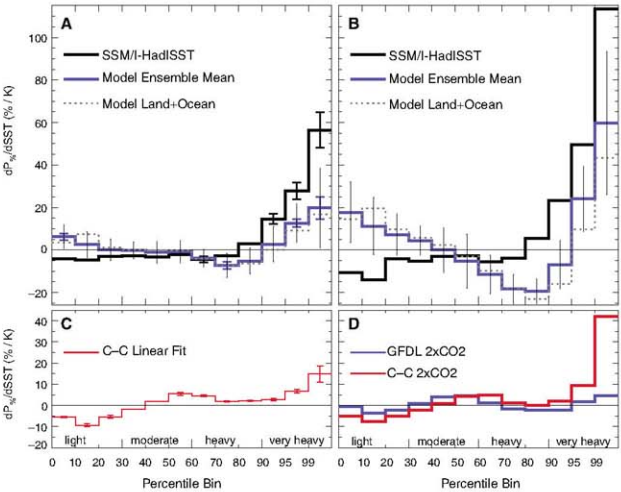


Fig. 4. Observed (black) and simulated (blue) percentage changes in precipitation frequency ($P_{\%}$) per K temperature increase over the tropical oceans calculated from **(A)** linear regressions, **(B)** El Niño minus La Niña, **(C)** the C-C experiment, and **(D)** the global warming response in the GFDL CM2.1–coupled model with A1B scenario forcings (*19*) using 2101–2105 minus 2001–2005 daily data. Also shown in (A) and (B) are the simulated changes for land and ocean (dotted line). Vertical lines in (A) and (B) denote ±1 standard deviation from all the individual models. Error bars in (A) and (C) represent ±1.96 standard errors for the linear fits plotted only where the correlation coefficient is greater in magnitude than 0.4 (the two-tailed test 95% confidence limit assuming 22 degrees of freedom). Also shown in (D) is the C-C response (red line; 7% times the local SST increase for 2101–2105 minus 2001–2005).

ble because moisture convergence in the tropics is itself determined by latent heat released via precipitation (*7*, *18*).

What implications do the differing precipitation responses in models and observations have for climate prediction? With the GFDL CM2.1 fully coupled climate model (*21*) scenario A1B (*19*), we considered the periods 2001–2005 and 2101–2105 (Fig. 4D). The simulated response shows an increased frequency of very heavy and moderate precipitation at the expense of light and heavy precipitation. Even accounting for the negative direct impact of CO_2 increases on precipitation (*7*, *24*) (SOM text), the response of the heaviest precipitation to warming is lower than that expected from C-C and compared with the present-day simulations (fig. S6). This raises the question of how climate model predictions of precipitation can be successfully evaluated. Nevertheless, the apparent underestimate of model-simulated response of the heaviest precipitation to SST changes compared with satellite data over the period 1987–2004 may indicate that climate projections also underestimate this response.

This study used natural climate variability to demonstrate a direct link between a warmer climate and an increase in extreme precipitation events in both satellite observations and model simulations. Although the models qualitatively reproduce the observed behavior, the rate of amplification of extreme rainfall events to atmospheric warming is found to be weaker in the models compared with observations. Similar discrepancies have been noted in global-mean trends in precipitation and evaporation (*11*, *15–17*). It also implies that model projections of future changes in extreme precipitation events in response to global warming may also be underpredicted. Of equal concern to water-depleted land regions are the responses of light and moderate precipitation on the periphery of convection (*1–4*, *25*), which also appear poorly captured by model simulations of the present-day climate. Given the potential social and economic implications of these findings, it is crucial to establish whether the discrepancy can be explained by inadequacies in the observing system (*26*), by the representation of decadal changes in aerosol-driven radiative forcing and associated surface flux changes (*27–29*), or by deficiencies in model parametrizations and simulation of current rainfall distributions (*20*).

References and Notes

1. G. Meehl *et al.*, in *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change* (Cambridge Univ. Press, Cambridge, 2007), pp. 747–845.
2. R. Seager *et al.*, *Science* **316**, 1181 (2007); published online 4 April 2007 (10.1126/science.1139601).
3. C. Chou, J. Tu, P. Tan, *Geophys. Res. Lett.* **34**, L17708 (2007).
4. J. D. Neelin, M. Munnich, H. Su, J. E. Meyerson, C. E. Holloway, *Proc. Natl. Acad. Sci. U.S.A.* **103**, 6110 (2006).
5. S. Emori, S. J. Brown, *Geophys. Res. Lett.* **32**, L17706 (2005).
6. G. A. Meehl, J. M. Arblaster, C. Tebaldi, *Geophys. Res. Lett.* **32**, L18719 (2005).
7. M. R. Allen, W. J. Ingram, *Nature* **419**, 224 (2002).
8. K. E. Trenberth, A. Dai, R. M. Rasmussen, D. B. Parsons, *Bull. Am. Met. Soc.* **84**, 1205 (2003).
9. I. M. Held, B. J. Soden, *J. Clim.* **19**, 5686 (2006).
10. B. J. Soden, D. L. Jackson, V. Ramaswamy, M. D. Schwarzkopf, X. Huang, *Science* **310**, 841 (2005); published online 6 October 2005 (10.1126/science.1115602).
11. F. J. Wentz, L. Ricciardulli, K. Hilburn, C. Mears, *Science* **317**, 233 (2007); published online 30 May 2007 (10.1126/science.1140746).
12. L. Bengtsson *et al.*, *Tellus* **59A**, 539 (2007).
13. R. P. Allan, *J. Geophys. Res.* **111**, D22105 (2006).
14. G. A. Vecchi *et al.*, *Nature* **441**, 73 (2006).
15. X. Zhang *et al.*, *Nature* **448**, 461 (2007).
16. R. P. Allan, B. J. Soden, *Geophys. Res. Lett.* **34**, L18705 (2007).
17. L. Yu, R. A. Weller, *Bull. Am. Met. Soc.* **88**, 527 (2007).
18. P. Pall, M. R. Allen, D. A. Stone, *Clim. Dyn.* **28**, 351 (2007).
19. G. A. Meehl *et al.*, *Bull. Am. Met. Soc.* **88**, 1383 (2007).
20. E. M. Wilcox, L. J. Donner, *J. Clim.* **20**, 53 (2007).
21. T. L. Delworth *et al.*, *J. Clim.* **19**, 643 (2006).
22. N. A. Rayner *et al.*, *J. Geophys. Res.* **108**, 4407 (2003).
23. K. E. Trenberth, A. Dai, *Geophys. Res. Lett.* **34**, L15702 (2007).
24. F. Yang, A. Kumar, M. E. Schlesinger, W. Wang, *J. Clim.* **16**, 2419 (2003).
25. C. E. Chung, V. Ramanathan, *Geophys. Res. Lett.* **34**, L16809 (2007).
26. F. R. Robertson, D. E. Fitzjarrald, C. D. Kummerow, *Geophys. Res. Lett.* **30**, 1180 (2003).
27. M. I. Mishchenko *et al.*, *Science* **315**, 1543 (2007).
28. M. Wild *et al.*, *Science* **308**, 847 (2005).
29. V. Ramanathan *et al.*, *Nature* **448**, 575 (2007).

30. Thanks to the World Climate Research Programme for enabling the Program for Climate Model Diagnosis and Intercomparison model archive (www-pcmdi.llnl.gov). R.A. was supported by the UK Natural Environment Research Council grants NE/C51785X/1 and the National Centre for Earth Observation. B.S. was supported by grants from National Oceanic and Atmospheric Administration Climate Prediction Office and NASA Energy and Water Cycle Study. Global Precipitation Climatology Project (GPCP) data were extracted from www.ncdc.noaa.gov; SSM/I data were provided by Remote Sensing Systems. Comments by K. Hodges and two anonymous reviewers helped to improve the manuscript.

Supporting Online Material

www.sciencemag.org/cgi/content/full/1160787/DC1
Materials and Methods
SOM Text
Figs. S1 to S7

21 May 2008; accepted 29 July 2008
Published online 7 August 2008;
10.1126/science.1160787
Include this information when citing this paper.

Superiority, Competition, and Opportunism in the Evolutionary Radiation of Dinosaurs

Stephen L. Brusatte,*† Michael J. Benton, Marcello Ruta, Graeme T. Lloyd

The rise and diversification of the dinosaurs in the Late Triassic, from 230 to 200 million years ago, is a classic example of an evolutionary radiation with supposed competitive replacement. A comparison of evolutionary rates and morphological disparity of basal dinosaurs and their chief "competitors," the crurotarsan archosaurs, shows that dinosaurs exhibited lower disparity and an indistinguishable rate of character evolution. The radiation of Triassic archosaurs as a whole is characterized by declining evolutionary rates and increasing disparity, suggesting a decoupling of character evolution from body plan variety. The results strongly suggest that historical contingency, rather than prolonged competition or general "superiority," was the primary factor in the rise of dinosaurs.

The rise of the dinosaurs in the Late Triassic and Early Jurassic (230 to 190 million years ago) is a classic example of an evolutionary radiation. During that time, the clade Dinosauria expanded from a single lineage to many dozens of lineages and from one ecological and morphological type to many, and the range of body sizes expanded to include truly gigantic forms (*1*, *2*). Through this expansion in diversity and disparity, dinosaurs became the preeminent vertebrates on land, occupying many ecological roles—especially those at medium to large size—in terrestrial ecosystems worldwide.

The expansion of Dinosauria has long been seen as an example of a "competitive adaptive radiation" in which one group supplants another (*3*, *4*). The dinosaurs were said to have outcompeted other terrestrial tetrapods (notably basal archosaurs, rhynchosaurs, and nonmammalian synapsids) by virtue of their upright or erect posture, which gave them advantages of speed and maneuverability (*5*), or because they were endothermic (possessing fully warm-blooded physiology) (*6*). The alternative, opportunistic model (*3*) proposes that dinosaurs diversified in the Norian, after a Carnian-Norian extinction event (CNEE) 228 million years ago that saw the demise of rhynchosaurs, dicynodonts, and chiniquodontids; dinosaurian clades were then added through the Late Triassic and Early Jurassic until they reached their full diversity.

Most previous studies have treated the rise of the dinosaurs as a single event, whether competitive or opportunistic (*3*, *5*–*8*). However, phylogenies and diversity trends suggest that this was a two-step process, with the diversification of herbivorous sauropodomorphs in the Norian (after the CNEE) followed by larger theropods and armored herbivore groups in the Early Jurassic [after extinction of carnivorous crurotarsans at or near the Triassic-Jurassic boundary (TJEE)]. This two-step model has been supported by recent studies of theropods, which became larger and more common after the TJEE (*7*), and ornithischians, which are now known to have been rare in the Late Triassic after the reassignment of many supposed ornithischian fossils to non-dinosaurian groups (*9*).

The critical interval to consider is the Late Triassic, especially the Norian and Rhaetian (Fig. 1), a 28-million-year span between the CNEE and TJEE. The key "competitors" of the early dinosaurs were the crurotarsans, the "crocodile-line" archosaurs, which show a range of morphologies and adaptations during this time: long-snouted fish- and flesh-eating phytosaurs, armored herbivorous aetosaurs, and large to giant carnivorous "rauisuchians." The crurotarsans even replicated many dinosaurian body plans (large terrestrial predators; small swift predators; mid- to large-bodied low-browsing herbivores; agile bipedal herbivores). Several new discoveries show striking convergences between crurotarsans and dinosaurs (*10*), and many Triassic crurotarsans were previously erroneously identified as dinosaur ancestors (*11*) or even as true dinosaurs (*12*). Such morphological convergence suggests that dinosaurs and crurotarsans were exploiting similar resources in the Late Triassic. In some Norian faunas, crurotarsans were numerically more abundant than dinosaurs (*3*) and seem to have exploited a wider range of body plans. However, by the end of the Triassic all crurotarsans were extinct, save a few lineages of crocodylomorphs.

The key question is why the major dinosaur lineages survived the TJEE, ushering in the 135-million-year "age of dinosaurs," while most crurotarsan groups went extinct. One common explanation is that dinosaurs outcompeted crurotarsans in the Late Triassic, and notions of general dinosaurian "superiority" have long pervaded the literature (*5*, *6*). Hypotheses of competition between major clades are often vague, difficult to test conclusively, and prone to oversimplification (*4*). Rather than focusing on such imprecise terms, it is illuminating to examine macroevolutionary patterns. Here, we compared evolutionary rates and relative morphospace occupation in dinosaurs and crurotarsans, in an effort to shed light on their evolutionary dynamics and to assess long-standing perceptions such as "superiority."

We used a new phylogeny of Triassic archosaurs (Fig. 1) and a data set consisting of 64 taxa and 437 discrete skeletal characters (*13*) to calculate numerical measures of evolutionary rates [patristic dissimilarity per branch and patristic dissimilarity divided by branch duration (*14*)] and disparity (morphospace occupation) (*13*). Note that evolutionary rates analysis approximates the amount of morphologic evolution separating species, whereas disparity analysis approximates the amount of morphologic difference between species (*14*–*17*). These are related but separate measures of morphological evolution that together give insights into patterns of macroevolutionary change within and between clades. Disparity analysis does not depend on a specific phylogenetic hypothesis, but evolutionary rates analysis does.

There is no clear evidence for differences in overall evolutionary rates between dinosaurs and crurotarsans during the Triassic as a whole. Dinosaurs exhibit higher mean rates than crurotarsans for all measures (Fig. 2, A and B, and fig. S2, A and B)—as does the entire dinosaur "total group," Ornithodira (sister taxon to Crurotarsi)—but these differences are generally not significant (tables S1 and S28). A pruned analysis of equal sample sizes for the two clades returns the same result (table S2), as does an analysis restricted to Norian taxa (table S3). There is limited evidence for significantly higher rates in Carnian dinosaurs, but this may be due to small sample size (table S3). Temporal trends do not show a coupled increase in dinosaur rates and decrease in crurotarsan rates, as might be expected under some models of "competition" (Fig. 2, E and F, and fig. S2, E and F). Relative to crurotarsans, dinosaurs exhibit a significantly higher rate of evolution of the appendicular skeleton, but not of the cranial or axial skeleton (table S9). However, there are no significant differences between rates for different regions of the dinosaur skeleton (tables S16 and S17).

Perhaps counterintuitively, the disparity study shows that crurotarsans occupied a larger amount of morphospace than did dinosaurs and ornithodirans as a whole (Fig. 3, A and B, fig. S3, A and B, and tables S21 and S29). Rarefaction curves show that these results are not biased by sample size (fig. S4). The same pattern holds within the Carnian and Norian (table S22), and there are no coupled temporal trends (Fig. 3, E and F, and fig. S3, E and F). Dinosaurs and crurotarsans occupy adjacent areas of morphospace (Fig. 1), which is expected because the analysis is based on cladistic characters.

Department of Earth Sciences, University of Bristol, Bristol BS8 1RJ, UK.

*Present address: Division of Paleontology, American Museum of Natural History, New York, NY 10024, USA, and Department of Earth and Environmental Sciences, Columbia University, Palisades, NY 10964, USA.

†To whom correspondence should be addressed. E-mail: brusatte@uchicago.edu

Crurotarsans convergent with dinosaurs (poposauroids, "rauisuchids," and basal crocodylomorphs) occupy an intermediate area between the majority of crurotarsans and dinosaurs. The higher disparity of crurotarsans is borne out by visual examination of Fig. 1, which shows a much larger morphospace than that for dinosaurs. Unexpectedly, this larger crurotarsan morphospace is associated with significantly higher rates of homoplasy (table S18), which suggests that character oscillation is an important factor in body plan evolution.

Archosaurs radiated during the Triassic in the aftermath of the end-Permian mass extinction. Our analysis shows that this radiation was associated with declining evolutionary rates per lineage and increasing morphological disparity throughout the Triassic. One rate metric—dissimilarity calibrated by time interval duration—shows a general decrease through the Triassic, with significantly high rates in the Anisian and low rates in the Norian (Fig. 2D, fig. S2D, and tables S4 and S5). Patterns within Crurotarsi and Dinosauria mirror those of Archosauria as a whole, as both subclades are characterized by decreasing rates (Fig. 2F, fig. S2F, and tables S6 to S8). Similarly, decreasing rates are also seen in cranial, axial, and appendicular character partitions (tables S10 to S15). The significantly high rates of character evolution in early archosaur history are consistent with the hypothesis of elevated rates during major morphological radiations (*15*, *18*–*20*).

In contrast, archosaurs show increasing disparity throughout the Triassic, with a significantly high peak in the Norian (Fig. 3, C and D, and tables S23 and S24). Both crurotarsans and dinosaurs show a general increase in disparity across the Triassic, except for a Ladinian drop for crurotarsans that may be due to small sample size, but the differences between time bins are not significant (tables S25 and S26). This pattern differs from the findings of several paleontological studies, which have shown that disparity often peaks early in the history of major clades (*15*–*17*, *21*).

Unexpectedly, these results indicate a decoupling of character evolution and morphological disparity in Triassic archosaurs (*22*, *23*). The inverse relationship indicates that, apparently, the burst of character evolution in early archosaur history did not translate into a wide range of body plans. Only later, when evolutionary rates decreased and homoplasy increased (tables S19 and S20), did a slower rate of character change result in the development of several new body plans (phytosaurs, aetosaurs, crocodylomorphs, pterosaurs, and dinosaurs), all of which are first known from the Carnian or Norian. Decoupling of lineage diversification and disparity has been noted before, but only in the context of within-subclade disparity among extant lizards (*24*). Further work is needed to determine what, if any, broad generalizations characterize evolutionary radiations across a wide range of organisms, time scales, and clade dimensions.

For the first 30 million years of their history, dinosaurs lived alongside and shared niches with another major clade (Crurotarsi) that occupied more morphospace and evolved at indistinguishable rates. These patterns seriously contrast with general notions of dinosaurian "superiority" and the long-standing view that dinosaurs were predordained for success (*5*, *6*). It is difficult to explain why crurotarsans and not dinosaurs went extinct at the TJEE, which may have been a catastrophic event (*7*) or an ecologically drawn-out affair triggered by eruption and elevated CO_2 levels (*25*). Either way, as in most mass extinction events, the death of species is often more random than ecologically selective (*26*), and so the relative proportions (or relative success) of two groups during normal times may reverse during a sudden crisis. Nonetheless, our results are consistent with at least two explanations: (i) Crurotarsans died out by chance, despite their larger range of morphospace and similar evolutionary rates to dinosaurs; (ii) dinosaurs prevailed because of one or several key adaptations. The second suggestion is difficult to entertain because dinosaurs and crurotarsans lived side by side for 30 million years, and crurotarsans occupied more morphospace and were often more abundant and diverse than dinosaurs. It is likely that dinosaurs were the beneficiaries of two mass extinction events—and some good luck.

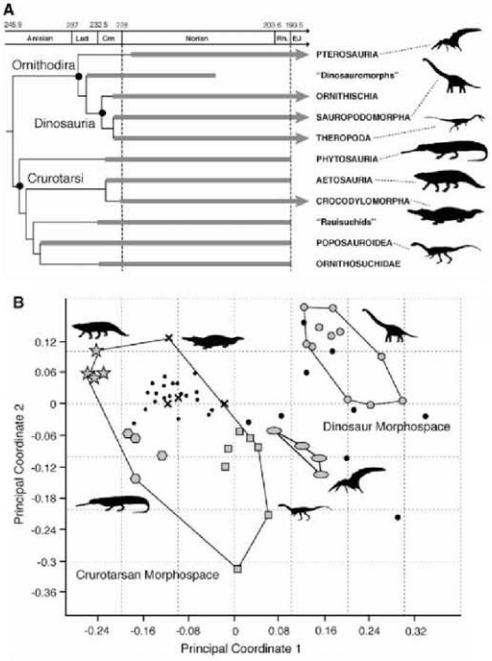


Fig. 1. Phylogenetic relationships and morphospace occupation for Triassic archosaurs. **(A)** Framework phylogeny for Triassic crurotarsans (*13*) scaled to the Triassic time scale (*13*). Numbers at top refer to millions of years before present; gray bars represent the observed durations of major lineages; vertical dashed lines denote two hypothesized extinction events (CNEE and TJEE); arrowheads indicate lineages that survived the TJEE. Lad, Ladinian; Crn, Carnian; Rh, Rhaetian; EJ, Early Jurassic. **(B)** Empirical morphospace for Triassic archosaurs, based on the first two principal coordinates (*13*). Large circles, dinosaurs; ovals, pterosaurs; squares, poposauroids; hexagons, phytosaurs; stars, aetosaurs; crosses, crocodylomorphs; smaller solid circles, "rauisuchids"; larger solid circles, nondinosaurian dinosauromorphs, *Scleromochlus*.

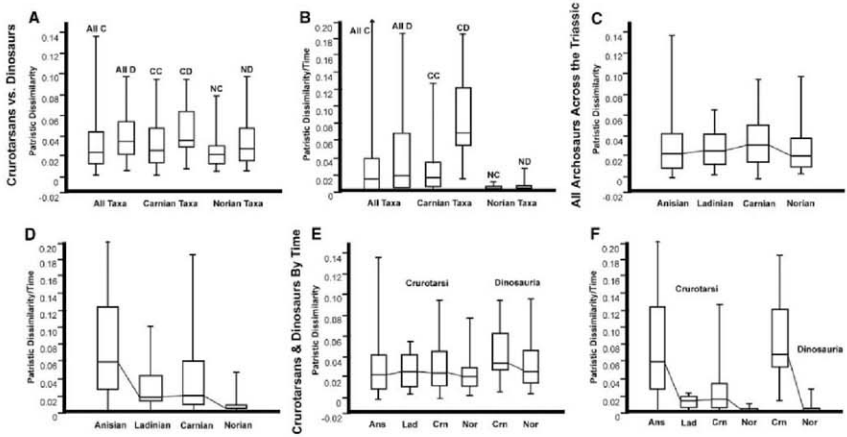


Fig. 2. Plots of rate of morphological character evolution for archosaurs, based on two metrics: patristic dissimilarity per branch and dissimilarity per time interval (*13*). Rates are based on ACCTRAN character optimization, but DELTRAN gives nearly identical results (fig. S2). Boxes represent the distribution of real data, with boxes encompassing percentiles 25 to 75 and the whiskers representing percentiles 5 to 95. (**A** and **B**) Evolutionary rates of crurotarsans and dinosaurs (All C, all Triassic crurotarsans; All D, all Triassic dinosaurs; CC, CD, NC, and ND, crurotarsans and dinosaurs subdivided into Carnian and Norian taxa). (**C** and **D**) Disparity against time for all crown-group archosaurs. (**E** and **F**) Disparity against time for both crurotarsans and dinosaurs. Dinosaurs exhibit higher evolutionary rates than crurotarsans, but these are not significant (table S1). Rates for all archosaurs are either approximately constant (dissimilarity metric) or decrease from an Anisian high to a Norian low (dissimilarity/time metric; see also tables S4 and S5). Patterns within Crurotarsi and Dinosauria mirror the general pattern (tables S6 to S9).

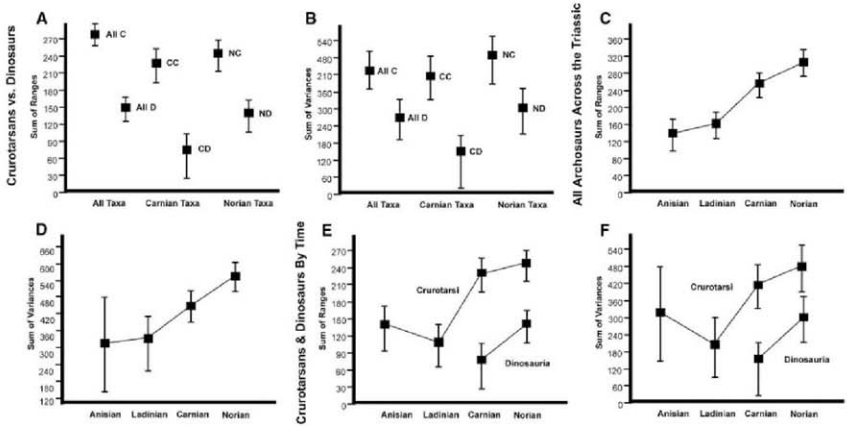


Fig. 3. Plots of archosaur morphological disparity, based on two metrics: sum of ranges and sum of variances (*13*). Squares represent mean values; error bars denote 95% confidence intervals based on bootstrapping. (**A** and **B**) Disparity of crurotarsans and dinosaurs (abbreviations as in Fig. 2). (**C** and **D**) Disparity against time for all crown-group archosaurs. (**E** and **F**) Disparity against time for both crurotarsans and dinosaurs. Crurotarsans exhibit a significantly higher disparity than dinosaurs when all Triassic taxa (nonparametric multivariate analysis of variance: $F = 29.89$, $P < 0.0001$) and Carnian ($F = 13.36$, $P = 0.0003$) and Norian ($F = 20.59$, $P < 0.0001$) subdivisions are analyzed. Archosaur disparity increases over time and reaches a statistically significant peak in the Norian (tables S23 and S24). Crurotarsan and dinosaur disparity generally increase over time, but differences between individual time bins are not significant (tables S25 and S26).

References and Notes

1. P. C. Sereno, *Science* **284**, 2137 (1999).
2. D. B. Weishampel, P. Dodson, H. Osmólska, Eds., *The Dinosauria* (Univ. of California Press, Berkeley, CA, ed. 2, 2004).
3. M. J. Benton, *Q. Rev. Biol.* **58**, 29 (1983).
4. M. J. Benton, *Biol. Rev. Cambridge Philos. Soc.* **62**, 305 (1987).
5. A. J. Charig, *Symp. Zool. Soc. London* **52**, 597 (1984).
6. R. T. Bakker, *Evolution* **25**, 636 (1971).
7. P. E. Olsen *et al.*, *Science* **296**, 1305 (2002).
8. R. B. Irmis *et al.*, *Science* **317**, 358 (2007).
9. S. J. Nesbitt, R. B. Irmis, W. G. Parker, *J. Syst. Palaeontol.* **5**, 209 (2007).
10. S. J. Nesbitt, M. A. Norell, *Proc. R. Soc. London Ser. B* **273**, 1045 (2006).
11. Examples include *Ornithosuchus* and *Postosuchus*.
12. Examples include *Paposaurus*, *Revueltosaurus*, *Shuvosaurus*, and *Teratosaurus*.
13. See supporting material on *Science* Online.
14. M. Ruta, P. J. Wagner, M. I. Coates, *Proc. R. Soc. London Ser. B* **273**, 2107 (2006).
15. P. J. Wagner, *Paleobiology* **23**, 115 (1997).
16. M. A. Wills, D. E. G. Briggs, R. A. Fortey, *Paleobiology* **20**, 93 (1994).
17. M. Foote, *Paleobiology* **20**, 320 (1994).
18. J. W. Valentine, *Paleobiology* **6**, 444 (1980).
19. D. Schluter, *The Ecology of Adaptive Radiation* (Oxford Univ. Press, Oxford, 2000).
20. S. J. Gould, *The Structure of Evolutionary Theory* (Harvard Univ. Press, Cambridge, MA, 2002).
21. D. H. Erwin, *Palaeontology* **50**, 57 (2007).
22. The results are unexpected under a null model of diffusive evolution with constant step size over time (*23*).
23. M. Foote, in *Evolutionary Paleobiology* (Univ. of Chicago Press, Chicago, 1996).
24. L. J. Harmon, J. A. Schulte II, A. Larson, J. B. Losos, *Science* **301**, 961 (2003).
25. J. C. McElwain, D. J. Beerling, F. I. Woodward, *Science* **285**, 1386 (1999).
26. D. Jablonski, *Science* **231**, 129 (1986).

27. We thank R. Benson, M. Coates, P. Donoghue, M. Foote, M. LaBarbera, and S. Wang for discussion. Supported by the Marshall Scholarship and Paleontological Society (S.L.B.), the UK Natural Environment Research Council (M.J.B., M.R., G.T.L.), and the Royal Society (M.J.B., M.R.).

Supporting Online Material

www.sciencemag.org/cgi/content/full/321/5895/1485/DC1
Materials and Methods
SOM Text
Figs. S1 to S4
Tables S1 to S29
References

16 June 2008; accepted 1 August 2008
10.1126/science.1161833

Niche Partitioning Increases Resource Exploitation by Diverse Communities

Deborah L. Finke[1,2]* and William E. Snyder[1]

Classical ecological theory suggests that the coexistence of consumer species is fostered by resource-use differences, leading to greater resource use in communities with more species. However, explicit empirical support for this idea is lacking, because resource use by species is generally confounded with other species-specific attributes. We overcame this obstacle by co-opting behavioral plasticity in food choice among a group of animal consumers, allowing us to manipulate patterns of resource use while controlling for the effects of species identity and diversity. Within an aphid-parasitoid-radish community, we created a fully factorial manipulation of consumer resource-use breadth (specialist versus generalist) and species diversity (one versus three species) and found that resource exploitation improved with greater specialist, but not generalist, diversity. Therefore, resource partitioning, and not diversity per se, fostered greater overall resource consumption in our multispecies consumer communities.

Early ecological models suggested that relatively strong intraspecific competition paired with relatively weak interspecific competition fosters species coexistence and promotes biodiversity (*1*–*4*). When these conditions exist, new species are able to invade model communities because they can monopolize a subset of the total resource pool. In contrast, when interspecific competition is the predominant force and resource partitioning is absent, only the single consumer species that drives the limiting resource to the lowest level is able to persist (*5*). This leads to the prediction that when species differ in resource-use patterns, adding more species to a community will lead to increased overall exploitation of available resources (*3*, *5*, *6*). It is resource differentiation among consumers at the community level that is expected to lead to more complete resource exploitation and not species diversity per se. However, empirical validation of these ideas has been hindered by the fact that resource-use differences among species typically are inextricably confounded with other species-specific attributes and requirements (such as size, rate of growth, metabolic rate, and fecundity). This lack of empirical support led, until recently, to the deemphasis of resource partitioning as a key driver of community structure (*1*).

Recent experimental manipulations of species richness have revealed, across a broad range of real-world ecological communities, a general pattern of greater resource exploitation when more species are present (*7*–*9*). However, the role of resource-use partitioning as a mechanism underlying this pattern, if any, has resisted empirical documentation (*10*–*16*). Progress has been hindered again by the seeming impossibility of entirely isolating the impacts of resource partitioning from those of other species attributes (*12*, *14*, *17*).

Here, we report an empirical test of the idea that resource partitioning leads to a net increase in resource exploitation by consumer communities. Our work was conducted in a model system in which plastic prey-choice behavior by natural enemies was exploited to manipulate overlap in resource use, independent of consumer species identity and thus of other species-specific traits. The system consisted of radish host plants, aphid herbivores, and parasitoid natural enemies. Radish (*Raphanus sativus*) plants in the Pacific Northwest of the United States are consumed by a variety of phloem-feeding aphid species, including green peach aphids (*Myzus persicae*), cabbage aphids (*Brevicoryne brassicae*), and turnip aphids (*Lipaphis erysimi*). These aphids are attacked by a diverse community of parasitoid wasps in the family Braconidae, including the species *Diaeretiella rapae*, *Aphidius colemani*, and *A. matricariae* (*18*). Insect parasitoids deliver natural pest control in agricultural systems worldwide, an ecosystem service of great economic and environmental value to humans (*19*).

We manipulated the resource use of individual consumer species by taking advantage of the natural host fidelity exhibited by these otherwise generalist parasitoid wasps (*18*, *20*). Although each parasitoid species is capable of attacking and completing development in all three aphid species, when given a choice, individual female wasps prefer to deposit eggs in hosts of the same species from which they themselves emerged (*20*) (fig. S1). This host fidelity is most likely expressed through associative learning. Upon emergence as adults, wasp parasitoids use the chemical cues associated with the natal host and its environment to direct their searching (*20*). As a result, parasitoids are more likely to locate and oviposit in hosts of the same species as their natal host. Such host fidelity behavior gave us an opportunity to manipulate the breadth of resources exploited by different populations of a single species and also across communities including several wasp species (*21*). We reared wasps of each of the three species on each of the three species of aphids, for a total of nine different wasp/aphid species associations. Then, by combining individual wasps from these source colonies, we could experimentally construct wasp communities differing in intraspecific and/or interspecific resource-niche breadth (fig. S2). By doing so, we were able to isolate the effects of competition on a well-defined resource, the aphid community, from the effects of other parasitoid species attributes.

Wasp communities were assembled that differed in all combinations of species identity, resource-use overlap ("specialists" that partition resources

[1]Department of Entomology, Washington State University, Pullman, WA 99164, USA. [2]Division of Plant Sciences, University of Missouri, Columbia, MO 65211, USA.

*To whom correspondence should be addressed. E-mail: finkedl@missouri.edu

versus "generalists" that completely overlap in their resource use), and the potential for intraspecific and/or interspecific competition (with a parasitoid species richness of one versus three) (*21*). We did this in field cages containing all three aphid species and measured the resulting impacts on the percentage of aphid parasitism and on aphid abundance. The manipulation of resource-use overlap and competitive interactions among parasitoids resulted in four parasitoid treatments: (i) a single specialist parasitoid species (36 individual parasitoids of the same species, all reared from the same aphid host); (ii) three specialist parasitoid species, each of which prefers to attack a different aphid host (12 individuals of each of the three parasitoid species, with each species reared on a different aphid host); (iii) a single generalist parasitoid species (36 parasitoids of the same species, with 12 individuals reared from each of the three aphid hosts); and (iv) three generalist parasitoid species that completely overlap in their resource use (12 individuals of each of the three parasitoid species, with 4 individuals of each parasitoid species reared from each of the three aphid species). Every possible parasitoid/host species combination was included within each treatment, and these compositions constituted replicates within that treatment (table S1). Including all parasitoid/host species combinations ensured that our results could not be unduly influenced by any single parasitoid species or by any parasitoid/aphid species pairing (*22*, *23*). Total parasitoid abundance at the time of release was held constant at 36 adult females (9 females/m^2) across all treatments. This experiment was conducted under real-world conditions in large field cages at the Washington State University Research Station in Othello, Washington.

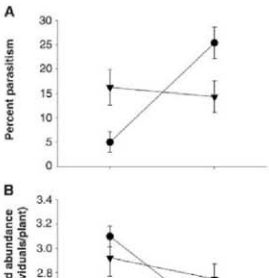


Fig. 1. Interactive effect of interspecific competitive interactions (with a parasitoid species richness of one versus three) and parasitoid resource-use differentiation on aphid population suppression. Population suppression by specialist parasitoids that partition their resource use (circles) is compared to that of generalist parasitoids with completely overlapping resource use (triangles) at two levels of species richness. At a species richness of one, only intraspecific interactions are possible, whereas both intraspecific and interspecific interactions are possible at a species richness of three. **(A)** Percent of total aphid population that is parasitized. **(B)** Total aphid abundance (log-transformed). Data are least squares means ± SEM obtained from repeated measures of analysis of covariance.

We found that parasitism success among wasp communities was affected by a strong interaction between the degree of resource-use overlap and consumer species richness (significant species richness times resource-use overlap interaction, $F_{1,32} = 12.56$, $P = 0.0013$; Fig. 1A). When parasitoids were generalists and any single species had access to all resources, increasing species richness did not affect the parasitism of the aphid community (*t* test of the difference between two means, $t_{32} = 0.40$, $P = 0.6934$; Fig. 1A). In contrast, when consumer species were specialists that used different resources, the percentage of parasitism increased dramatically when three species were present as opposed to one ($t_{32} = 5.25$, $P < 0.0001$; Fig. 1A). Comparing the two treatments including multiple consumer species, the percentage of parasitism was significantly greater when consumer species were specialists than generalists ($t_{32} = 2.40$, $P = 0.0224$; Fig. 1A). Aphid densities did not differ among treatments during the early course of the experiment (fig. S3), suggesting that parasitism rates were not indirectly affected by confounding differences in resource abundance among treatments.

Differences in the percentage of parasitism across treatments resulted in concordant differences in aphid densities. Parasitoid species richness and resource-use overlap interacted to determine total aphid abundance (significant species richness times resource-use overlap interaction, $F_{1,32} = 3.98$, $P = 0.0550$; Fig. 1B). Suppression of aphids was unaffected by the presence of multiple consumer species when parasitoids were generalists that completely overlap in their resource use ($t_{32} = 0.90$, $P = 0.3765$; Fig. 1B), suggesting equitability in the magnitude of intraspecific and interspecific interactions. Such competition among parasitoids is often chemically mediated, with parasitoid females being capable of recognizing the presence of both intraspecific and interspecific competitors (*24*). However, for specialist parasitoids, aphid consumption was greater and thus aphid abundance was lower, with greater parasitoid species richness ($t_{32} = 4.40$, $P = 0.0007$; Fig. 1B). Consistent with these results, per capita impacts on aphids of specialist parasitoids but not generalist parasitoids were higher with greater parasitoid species richness (fig. S4).

We independently manipulated resource-niche breadth and consumer species richness and found that resource exploitation was strengthened by a complex interaction between these two factors. Among our treatment combinations, the most substantial parasitism of aphids, and thus the lowest aphid densities, were recorded in communities combining multiple species of specialist parasitoids. In contrast, wasp performance was relatively weaker in diverse communities of generalists. With species richness held constant, the key difference between these two treatments is that we would expect intraspecific competition to be relatively intense and interspecific competition relatively weak for diverse communities of specialists as compared to generalists (*25*). Thus, our results closely match the preconditions for species coexistence predicted by classic early niche models (*2*, *3*). Additionally, our results match more recent assertions that it is differences in resource use among species, rather than diversity per se, that intensifies resource exploitation at higher levels of consumer diversity (*6*, *16*, *26*–*28*). Thus, we found empirical evidence that resource-niche partitioning may be both a factor encouraging greater biodiversity and an underlying cause of efficient resource extraction by species-rich communities, once assembled. Our results also support the argument that it is the conservation of species that fulfil specialized functional roles, rather that greater diversity itself, that is needed to preserve ecosystem function (*14*, *29*, *30*).

Studies focusing on predaceous animal consumers can be particularly enlightening, because resources (prey) in such systems are easily identified and the effects of resource capture (prey suppression) are readily observable (*13*, *16*, *25*, *31*, *32*). Further, when foraging behavior is plastic, differences in resource use among species can be experimentally manipulated, a powerful technique for testing the predictions of theoretical models related to resource partitioning, species coexistence, and biodiversity.

References and Notes

1. J. M. Chase, M. A. Leibold, *Ecological Niches: Linking Classical and Contemporary Approaches* (Univ. of Chicago Press, Chicago, 2003).
2. G. E. Hutchinson, *Am. Nat.* **93**, 145 (1959).
3. R. MacArthur, *Theor. Popul. Biol.* **1**, 1 (1970).
4. R. B. McKane *et al.*, *Nature* **415**, 68 (2002).
5. P. Chesson, *Annu. Rev. Ecol. Syst.* **31**, 343 (2000).
6. R. M. May, *Stability and Complexity in Model Ecosystems* (Princeton Univ. Press, Princeton, NJ, 1973).
7. P. Balvanera *et al.*, *Ecol. Lett.* **9**, 1146 (2006).
8. B. J. Cardinale *et al.*, *Nature* **443**, 989 (2006).
9. B. Worm *et al.*, *Science* **314**, 787 (2006).
10. M. E. S. Bracken, J. J. Stachowicz, *Ecology* **87**, 2397 (2006).
11. J. D. Fridley, *J. Ecol.* **91**, 396 (2003).
12. A. Kahmen, C. Renker, S. B. Unsicker, N. Buchmann, *Ecology* **87**, 1244 (2006).
13. W. E. Snyder, G. B. Snyder, D. L. Finke, C. S. Straub, *Ecol. Lett.* **9**, 789 (2006).
14. D. Tilman *et al.*, *Science* **277**, 1300 (1997).
15. A. V. Tiunov, S. Scheu, *Ecol. Lett.* **8**, 618 (2005).
16. A. Wilby, M. B. Thomas, *Ecol. Lett.* **5**, 353 (2002).
17. H. R. Lambers, W. S. Harpole, D. Tilman, J. M. H. Knops, P. Reich, *Ecol. Lett.* **7**, 661 (2004).
18. K. Pike *et al.*, *Proc. Entomol. Soc. Wash.* **102**, 688 (2000).
19. J. E. Losey, M. Vaughan, *Bioscience* **56**, 311 (2006).
20. T. C. J. Turlings, F. L. Wäckers, L. E. M. Vet, W. J. Lewis, J. H. Tumlinson, in *Insect Learning: Ecological and Evolutionary Perspectives*, D. R. Papaj, A. C. Lewis, Eds. (Chapman & Hall, New York, 1992), pp. 51–78.
21. See supporting material on *Science* Online.
22. M. A. Huston, *Oecologia* **110**, 449 (1997).

23. D. Tilman, C. L. Lehman, K. T. Thomson, *Proc. Natl. Acad. Sci. U.S.A.* **94**, 1857 (1997).
24. C. R. Nufio, D. R. Papaj, *Entomol. Exp. Appl.* **99**, 273 (2001).
25. D. L. Finke, R. F. Denno, *Nature* **429**, 407 (2004).
26. P. Casula, A. Wilby, M. B. Thomas, *Ecol. Lett.* **9**, 995 (2006).
27. A. R. Ives, B. J. Cardinale, W. E. Snyder, *Ecol. Lett.* **8**, 102 (2005).
28. M. Loreau, *Proc. Natl. Acad. Sci. U.S.A.* **95**, 5632 (1998).
29. D. R. Chalcraft, W. J. J. Resetarits, *Ecology* **84**, 2407 (2003).
30. D. U. Hooper, J. S. Dukes, *Ecol. Lett.* **7**, 95 (2004).
31. T. Bukovinszky, F. J. F. van Veen, Y. Jongema, M. Dicke, *Science* **319**, 804 (2008).
32. O. J. Schmitz, *Science* **319**, 952 (2008).
33. We thank A. Ives, J. Owen, A. Storfer, C. Straub, and anonymous reviewers for comments. This project was supported by the National Research Initiative of the U.S. Department of Agriculture Cooperative Research, Education and Extension Service, grant 2004-01215.

Supporting Online Material
www.sciencemag.org/cgi/content/full/321/5895/1488/DC1
Materials and Methods
Figs. S1 to S4
Table S1
References

22 May 2008; accepted 11 August 2008
10.1126/science.1160854

Degradation of microRNAs by a Family of Exoribonucleases in *Arabidopsis*

Vanitharani Ramachandran and Xuemei Chen*

microRNAs (miRNAs) play crucial roles in numerous developmental and metabolic processes in plants and animals. The steady-state levels of miRNAs need to be properly controlled to ensure normal development. Whereas the framework of miRNA biogenesis is established, factors involved in miRNA degradation remain unknown. Here, we show that a family of exoribonucleases encoded by the *SMALL RNA DEGRADING NUCLEASE* (*SDN*) genes degrades mature miRNAs in *Arabidopsis*. SDN1 acts specifically on single-stranded miRNAs in vitro and is sensitive to the 2′-*O*-methyl modification on the 3′ terminal ribose of miRNAs. Simultaneous knockdown of three *SDN* genes in vivo results in elevated miRNA levels and pleiotropic developmental defects. Therefore, we have uncovered the enzymes that degrade miRNAs and demonstrated that miRNA turnover is crucial for plant development.

Plant miRNAs carry a 2′-*O*-methyl group that protects them from a 3′-to-5′ exonucleolytic activity and a uridylation activity that adds an oligoU tail to the 3′ ends of miRNAs (*1*, *2*). Maintaining proper steady-state levels of miRNAs is crucial for plant development (*3*–*7*). The steady-state levels of miRNAs are presumably determined by the opposing activities of miRNA biogenesis and degradation. A conserved exonuclease from *Caenorhabditis elegans* and *Schizosaccharomyces pombe*, Eri-1, specifically degrades small interfering RNA (siRNA)/siRNA* (where siRNA* represents antisense siRNA) duplexes with 2-nucleotide (nt) 3′ overhangs in vitro and reduces RNA interference efficiency in vivo (*8*, *9*). Exonucleases that degrade single-stranded small RNAs have yet to be identified.

To identify enzymes that degrade single-stranded miRNAs or siRNAs, we took a candidate-gene approach. We presume that enzymes involved in miRNA metabolism evolved from enzymes that process structural and/or catalytic RNAs, a view supported by the fact that a number of known players in small RNA metabolism also function in the processing of ribosomal RNAs (rRNAs) (*10*–*13*). We sought for *Arabidopsis* homologs of a class of exoribonucleases in yeast, Rex1p to Rex4p, which participate in 3′-end processing of rRNAs and tRNAs (*14*, *15*). BLAST (*16*) searches using the 4 Rex proteins identified 15 *Arabidopsis* proteins containing an exonuclease domain (fig. S1). At3g15140, which belongs to a clade of 6 proteins (fig. S1), was the most similar to Eri-1 among the 15 proteins. Because we seek enzymes that degrade single-stranded small RNAs, we excluded proteins in this clade from our analysis.

From the remaining Rex homologs, we randomly chose At3g50100 from the five-member clade and At3g15080 from the outliers (fig. S1), expressed them as glutathione *S*-transferase (GST) fusion proteins in *Escherichia coli* (fig. S2), and tested their activities on miRNAs in vitro (*17*). A 5′ end–labeled single-stranded RNA oligonucleotide corresponding to miR167 in sequence (but lacking a 2′-*O*-methyl group) was incubated with GST-At3g15080, GST-At3g50100, or GST. Whereas GST-At3g15080 or GST did not exhibit any activity on miR167, GST-At3g50100 degraded the full-length miR167, generating a product of ~8 to 9 nt (Fig. 1A; the size of the final product was estimated from Fig. 2D). GST-At3g50100 also acted on miR173 and 2′-*O*-methylated miR173 and generated products of ~8 to 9 nt (Fig. 1A). We refer to At3g50100 as SMALL RNA DEGRADING NUCLEASE1 (SDN1) hereafter.

To determine whether SDN1 is an endonuclease cleaving the RNAs between nucleotides 8 and 9 from their 5′ ends or a 3′-to-5′ exonuclease that cannot process RNAs of 8 nt or shorter, we labeled miR173 with ^{32}pCp at the 3′ end and incubated miR173-^{32}pCp with GST-SDN1. miR173-^{32}pCp was resistant to GST-SDN1, and phosphatase treatment of miR173-^{32}pCp to remove the 3′ phosphate rendered the miRNA susceptible to GST-SDN1 (Fig. 1B). Furthermore, a product of 15 nt, which would be expected if SDN1 were an endonuclease cleaving between nucleotides 8 and 9 from the 5′ end, was not observed on phosphatase-treated miR173-^{32}pCp (Fig. 1B). These data indicated that SDN1 is a 3′-to-5′ exonuclease.

GST-SDN1 did not have any effect on a single-stranded DNA oligonucleotide (Fig. 2B) and is therefore a ribonuclease. Unlike Eri-1 (*9*), GST-SDN1 failed to degrade miR173 in a miR173/miR173* duplex (Fig. 2B and fig. S3). To examine SDN1 substrate size, synthetic RNA oligonucleotides of 17, 18, 20, 21 (miR167), 22 (miR173), 23, 24, and 27 nt (table S2) were incubated with GST-SDN1 separately. SDN1 degraded all tested RNA oligonucleotides and yielded an end product of ~8 to 9 nt, regardless of the length of the substrates (Fig. 2A). However, SDN1 cannot act on longer RNAs. pre-miR167 or a 300-nt RNA from the protein-coding *APETALA1* (*AP1*) gene was not detectably degraded by GST-SDN1 (Fig. 2C). Therefore, SDN1 acts specifically on single-stranded small RNAs in a sequence-independent manner.

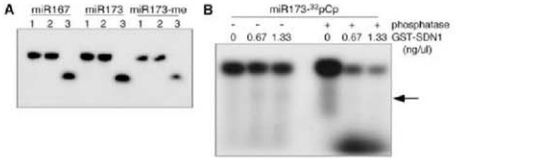


Fig. 1. *Arabidopsis* At3g50100 (SDN1) possesses 3′-to-5′ exonuclease activity on miRNAs. **(A)** Enzymatic activity assays on single-stranded miRNAs in vitro. RNA oligonucleotides were 5′-end labeled, incubated with buffer alone (1), purified GST (2), or purified GST-At3g50100 (3), and resolved on a denaturing polyacrylamide gel. miR173-me is a miR173 oligonucleotide containing a 2′-*O*-methyl group on the 3′ terminal ribose. **(B)** Enzymatic activity of GST-At3g50100 (GST-SDN1) on miR173 labeled at the 3′ end with ^{32}pCp. miR173-^{32}pCp was treated (+) or not treated (–) with phosphatase before incubation with GST-SDN1. The arrow indicates the position of the expected 15-nt product if SDN1 were to cleave the RNA between nucleotides 8 and 9 from the 5′ end. The radioactivity at the bottom corresponds to the position of free nucleotides.

Department of Botany and Plant Sciences, Institute of Integrative Genome Biology, University of California Riverside, Riverside, CA 92521, USA.

*To whom correspondence should be addressed. E-mail: xuemei.chen@ucr.edu

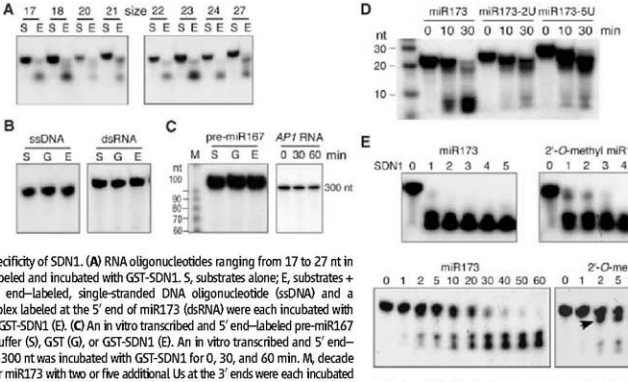


Fig. 2. Substrate specificity of SDN1. **(A)** RNA oligonucleotides ranging from 17 to 27 nt in length were 5′-end labeled and incubated with GST-SDN1. S, substrates alone; E, substrates + GST-SDN1. **(B)** A 5′ end–labeled, single-stranded DNA oligonucleotide (ssDNA) and a miR173/miR173* duplex labeled at the 5′ end of miR173 (dsRNA) were each incubated with buffer (S), GST (G), or GST-SDN1 (E). **(C)** An in vitro transcribed and 5′ end–labeled pre-miR167 was incubated with buffer (S), GST (G), or GST-SDN1 (E). An in vitro transcribed and 5′ end–labeled *AP1* RNA of ~300 nt was incubated with GST-SDN1 for 0, 30, and 60 min. M, decade marker. **(D)** miR173 or miR173 with two or five additional Us at the 3′ ends were each incubated with GST-SDN1 for 0, 10, or 30 min. **(E)** Effects of the 2′-*O*-methyl group on GST-SDN1 activity. (Top) 5′ end–labeled miR173 or 2′-*O*-methyl miR173 was incubated with increasing amounts of GST-SDN1 (numbers 0 to 5 represent 0 ng/μl, 0.33 ng/μl, 0.67 ng/μl, 1.33 ng/μl, 2.0 ng/μl, and 2.67 ng/μl of GST-SDN1, respectively). (Bottom) miR173 or 2′-*O*-methyl miR173 was incubated with 0.67 ng/μl GST-SDN1 for the specified time periods (min). The arrowhead indicates a ~20-nt intermediate. Under low enzyme concentration, another intermediate of 9 to 10 nt was also present [in (D) and (E)]. The bottom band corresponds to the final 8-to-9–nt product.

The 2′-*O*-methyl group present in all plant small RNAs (*1*, *2*) deters the activities of SDN1. When miR173 or 2′-*O*-methyl miR173 was incubated with varying concentrations of GST-SDN1, a degradation intermediate of ~20 nt was present in reactions on 2′-*O*-methyl miR173 under lower enzyme concentrations but was barely detectable in reactions on miR173 (Fig. 2E, top). In a time course using a low enzyme concentration (Fig. 2E, bottom), the rate of degradation of miR173 was faster than that of 2′-*O*-methyl miR173, as judged by the time of appearance of the final product. The 20-nt intermediate was much more prominent and lingered longer in the 2′-*O*-methyl miR173 reaction (Fig. 2E, bottom).

SDN1 is a multiple-turnover enzyme. In the reactions in Fig. 2E, the great majority of the substrates (4 pmol) was degraded by GST-SDN1 (278 fmol) in 60 min. Therefore, 1 molecule of enzyme degrades 14 molecules of small RNA.

miRNAs are uridylated on their 3′ ends when not methylated (*1*). miR173 with two or five additional Us on the 3′ end was not degraded as efficiently as was miR173 by GST-SDN1 (Fig. 2D), as judged by the delayed appearance of the final product and delayed disappearance of the full-length substrates or shorter intermediates. This suggests that uridylation of miRNAs in the absence of methylation could have a protective role against exonucleolytic degradation.

To determine whether SDN1 limits miRNA accumulation in vivo, we identified a homozygous transferred DNA (T-DNA) insertion mutant, *sdn1-1* (fig. S4). This mutant is not likely to be a null allele (fig. S5A), and it shows no obvious developmental defects or much difference in the abundance of seven tested miRNAs from that of the wild type (Fig. 3). The lack of miRNA defects in *sdn1-1* could be due to redundancy with the other four members of the clade: At3g50090, At5g05540 (*SDN2*), At5g67240 (*SDN3*), and At5g25800 (fig. S1). We obtained T-DNA insertion alleles in the three genes most closely related to *SDN1* (fig. S4). The abundance of seven tested miRNAs was largely unaffected in all four single mutants [*sdn1-1*; *sdn2-1*, probably a reduction-of-function allele (fig. S5); *sdn3-1*, a reduction-of-function or null allele (fig. S5); and the T-DNA allele in At3g50090, a possible pseudo gene] (Fig. 3). Three of the seven tested miRNAs (miR159, miR167, and miR173) and siR1003, an endogenous siRNA, accumulated to 1.5 to 1.8 times the wild-type (WT) levels in the *sdn1-1 sdn2-1* double mutant (Fig. 3).

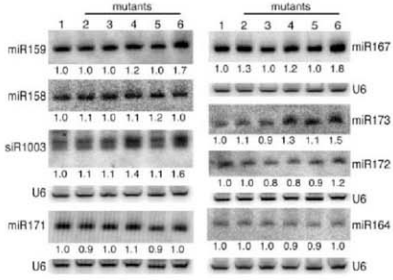


Fig. 3. Northern blot to detect the steady-state levels of seven miRNAs and an siRNA in mutants of *SDN1* and related genes. The U6 blots serve as a loading control. The numbers below the blots indicate the relative abundance of the small RNAs in the different genotypes. 1, wild type; 2, *sdn1-1*; 3, the mutant in At3g50090; 4, *sdn2-1*; 5, *sdn3-1*; and 6, *sdn1-1 sdn2-1*.

To further interrogate the gene family, we introduced an artificial miRNA (amiRNA) (*18*) that targets the exonuclease region in four of the five genes in the clade (fig. S4) into *sdn1-1*. In the T1 population, plants with various pleiotropic developmental defects were observed (Fig. 4, A to F, and table S3). Type I plants (Fig. 4, B to D), which were most severely affected, had small and often serrated leaves. Some plants had pin-like protrusions emanating from the abaxial side of the rosette leaves (Fig. 4C). Similar protrusions have been found in leaves of plants carrying an antisense *AGO1* cDNA or those undergoing sense *AGO1* cDNA-mediated cosuppression (*19*).

Levels of amiRNA-targeted *SDN1*, *SDN2*, and *SDN3* transcripts were severely reduced in one individual line and moderately reduced in another line (Fig. 4, H and I). miR167 accumulated to two to four times that of the WT level in the two amiRNA lines (Fig. 4G). Consistent with the presence of pinlike structures in the first individual line, a strong reduction in *AGO1* mRNA levels was found (Fig. 4J). To analyze the amiRNA lines more extensively, we pooled T1 plants according to the severity of the developmental phenotypes. Type I plants (lanes 1, 2, and 4 in Fig. 4, J and K) had the highest levels of the amiRNA (Fig. 4J), greatly reduced levels of

SDN1 and *SDN2* transcripts, and a slight reduction in *SDN3* transcript levels (Fig. 4K and fig. S6). In these lines, miR167, miR159, and siR1003 levels were to two to three times that of the wild type, and miR172 levels, which were not elevated in the *sdn1-1 sdn2-1* mutant (Fig. 3), were up to threefold of the WT level (Fig. 4J). The remaining miRNAs (except for miR164) all showed some elevation in abundance in some of the type I plants. Type II and III plants (lanes 3 and 5, respectively, in Fig. 4, J and K) had moderate levels of the amiRNA (Fig. 4J), a moderate-to-severe reduction in *SDN1* and *SDN2* transcript levels (Fig. 4K and fig. S6), and a moderate or no elevation in the abundance of endogenous small RNAs (Fig. 4J).

We did not observe any 3′ extended forms of the 5S or 5.8S rRNAs, which readily accumulate in the yeast *rex* mutants (*15*) and in the *C. elegans eri-1* mutant (*13*), respectively, in any of the *sdn* single mutants, the *sdn1-1 sdn2-1* double mutant, or the amiRNA lines (fig. S7). This result, together with the inability of SDN1 to digest small RNA duplexes, pre-miRNAs, or longer RNAs in vitro, suggests that single-stranded small RNAs are the most likely in vivo substrates of SDN1. However, a role for these genes in the metabolism of other classes of RNAs cannot be excluded.

In conclusion, we have identified a family of exonucleases that degrades single-stranded small RNAs in vitro and limits the accumulation of small RNAs in vivo. SDN1 and the only other known small RNA exonuclease, Eri-1, have distinct substrate specificities. The pleiotropic developmental phenotypes associated with reduction-of-function of the *SDN* gene family indicates that small RNA turnover is crucial for developmental patterning in plants. This family of genes is universally present in eukaryotes, and it is likely that the animal homologs of *SDN*1 perform similar functions in small RNA metabolism.

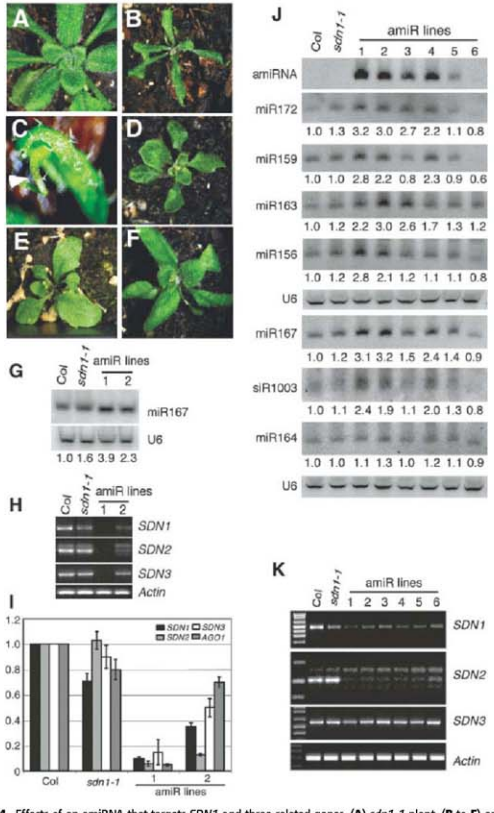


Fig. 4. Effects of an amiRNA that targets *SDN1* and three related genes. **(A)** *sdn1-1* plant. **(B** to **F)** amiRNA lines (in *sdn1-1*) with developmental defects of varying severity. (B to D) Type I plants. The arrowhead in (C) indicates a pinlike protrusion. (E) Type II plant that has small, mildly serrated leaves. (F) Early flowering type III plant. **(G)** Accumulation of miR167 in the wild type (Col), *sdn1-1*, and two individual type I amiRNA lines. The numbers below the blots indicate the relative abundance of the miRNA. **(H** and **I)** Levels of *SDN* and *AGO1* mRNAs, as determined by reverse transcription polymerase chain reaction (RT-PCR) (H) and real-time PCR (I), in the four genotypes shown in (G). Error bars in (I) indicate SD among three replicates. **(J)** Northern blotting to detect the amiRNA and endogenous small RNAs in pooled amiRNA lines. The numbers below the blots indicate the relative abundance of the small RNAs. **(K)** RT-PCR to detect *SDN* transcripts in the wild type (Col), *sdn1-1*, and six pools of amiRNA lines. The "–RT" controls did not yield any products and are not shown. The three bands for *SDN2* probably represent alternative transcripts because all three were missing in the –RT control.

References and Notes

1. J. Li, Z. Yang, B. Yu, J. Liu, X. Chen, *Curr. Biol.* **15**, 1501 (2005).
2. B. Yu *et al.*, *Science* **307**, 932 (2005).
3. M. J. Aukerman, H. Sakai, *Plant Cell* **15**, 2730 (2003).
4. X. Chen, *Science* **303**, 2022 (2004), published online 31 July 2003; 10.1126/science.1088060.
5. J. F. Palatnik *et al.*, *Nature* **425**, 257 (2003).
6. W. Park, J. Li, R. Song, J. Messing, X. Chen, *Curr. Biol.* **12**, 1484 (2002).
7. B. J. Reinhart, E. G. Weinstein, M. W. Rhoades, B. Bartel, D. P. Bartel, *Genes Dev.* **16**, 1616 (2002).
8. T. Iida, R. Kawaguchi, J. Nakayama, *Curr. Biol.* **16**, 1459 (2006).
9. S. Kennedy, D. Wang, G. Ruvkun, *Nature* **427**, 645 (2004).
10. H. Wu, H. Xu, L. J. Miraglia, S. T. Crooke, *J. Biol. Chem.* **275**, 36957 (2000).
11. T. Fukuda *et al.*, *Nat. Cell Biol.* **9**, 604 (2007).
12. K. M. Ansel *et al.*, *Nat. Struct. Mol. Biol.* **15**, 523 (2008).
13. H. W. Gabel, G. Ruvkun, *Nat. Struct. Mol. Biol.* **15**, 531 (2008).
14. A. W. Faber *et al.*, *RNA* **10**, 1946 (2004).
15. A. van Hoof, P. Lennertz, R. Parker, *EMBO J.* **19**, 1357 (2000).
16. The BLAST program, http://blast.ncbi.nlm.nih.gov/Blast.cgi.
17. Materials and methods are available as supporting material on *Science* Online.
18. R. Schwab, S. Ossowski, M. Riester, N. Warthmann, D. Weigel, *Plant Cell* **18**, 1121 (2006).
19. K. Bohmert *et al.*, *EMBO J.* **17**, 170 (1998).

20. We thank L. Bi for technical assistance and T. Dinh, L. Ji, B. Yu, and B. Zheng for helpful discussions and careful reading of the manuscript. This work was supported by grants from NSF (MCB-0718029) and NIH (GM61146) to X.C.

Supporting Online Material

www.sciencemag.org/cgi/content/full/321/5895/1490/DC1
Materials and Methods
Figs. S1 to S7
Tables S1 to S3
References

24 July 2008; accepted 11 August 2008
10.1126/science.1163728

Activation of Aldehyde Dehydrogenase-2 Reduces Ischemic Damage to the Heart

Che-Hong Chen,[1] Grant R. Budas,[1] Eric N. Churchill,[1] Marie-Hélène Disatnik,[1] Thomas D. Hurley,[2] Daria Mochly-Rosen[1]*

There is substantial interest in the development of drugs that limit the extent of ischemia-induced cardiac damage caused by myocardial infarction or by certain surgical procedures. Here, using an unbiased proteomic search, we identified mitochondrial aldehyde dehydrogenase 2 (ALDH2) as an enzyme whose activation correlates with reduced ischemic heart damage in rodent models. A high-throughput screen yielded a small-molecule activator of ALDH2 (Alda-1) that, when administered to rats before an ischemic event, reduced infarct size by 60%, most likely through its inhibitory effect on the formation of cytotoxic aldehydes. In vitro, Alda-1 was a particularly effective activator of ALDH2*2, an inactive mutant form of the enzyme that is found in 40% of East Asian populations. Thus, pharmacologic enhancement of ALDH2 activity may be useful for patients with wild-type or mutant ALDH2 who are subjected to cardiac ischemia, such as during coronary bypass surgery.

Cardiac ischemia is the leading cause of death. The discovery of a cardioprotective mechanism called preconditioning (induced by repetitive sublethal ischemic events) has triggered the search for pharmacological agents that mimic this effect (*1*, *2*). Adenosine (*3*), ethanol (*4*), and selective activation of protein kinase Cε (PKCε) (*4*, *5*) mimic ischemic preconditioning and reduce cardiac infarct size. Systematic searches for mediators of cardiac protection have identified a number of proteins whose levels or phosphorylation changes with cardioprotection (*6*, *7*). However, whether the changes were critical for cardiac protection was not determined.

We used an unbiased proteomic approach in ischemic rat hearts treated with ethanol and a selective inhibitor and an activator of PKCε that we generated (*4*, *7*, *8*). We found that one protein whose phosphorylation status consistently correlated with cardioprotection was mitochondrial aldehyde dehydrogenase 2 (ALDH2) (Fig. 1 and fig. S1, A to D). Under normoxic conditions, ALDH2 appeared as four phosphoproteins after isoelectric focusing (IEF) for two-dimensional (2-D) SDS gel electrophoresis [note 1 in (*5*)]. After preconditioning by a brief exposure to ethanol (50 mM, 10 min) (*9*) or selective activation of PKCε by the isozyme-specific agonist peptide, ψεRACK (receptor for activated C kinase) (1 μM, 10 min), which causes cardioprotection (*8*), there were only two (the more acidic) ALDH2 spots (Fig. 1A). The ethanol-induced shift in ALDH2 mobility was inhibited in the presence of the PKCε-selective antagonist peptide (εV1-2) (Fig. 1A), a treatment that we previously found to inhibit ethanol-induced cardiac protection (*9*). Therefore, ethanol-induced ALDH2 phosphorylation, which correlates with cardiac protection from ischemia, is dependent on PKCε activation.

How this mitochondrial enzyme was regulated by the cytosolic PKCε was not obvious. We first demonstrated that PKCε phosphorylates ALDH2 in vitro and that this phosphorylation results in a 38 ± 9% increase in ALDH2 catalytic activity [$n = 6$, $P < 0.005$ (fig. S2, A and B); notes 2 and 3 in (*5*)]. At least two phosphorylation sites were identified by mass spectroscopy, including Thr^{185} and Thr^{412} and possibly Ser^{279} [note 3 in (*5*)]. Further, coimmunoprecipitation of extracts from normoxic and ischemic hearts with antibodies against PKCε or ALDH2 confirmed the association of ALDH2 and PKCε in the mitochondrial fraction [fig. S3 and (*5*)]. Other subfractionation and biochemical studies have shown that PKCε regulates intramitochondrial proteins [e.g., (*6*)]. It is therefore likely that PKCε can enter

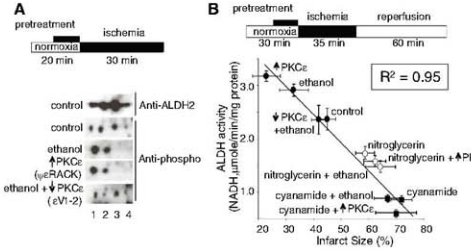


Fig. 1. (A) Ethanol and PKCε activation induce phosphorylation of mitochondrial ALDH2. Homogenates of rat hearts subjected to ischemia ex vivo were separated by IEF/SDS 2-D gel electrophoresis and probed with a mixture of phospho-serine and phospho-threonine antibodies. Using a Langendorff apparatus, hearts were perfused with oxygenated Krebs-Henseleit buffer alone as control, with 50 mM ethanol for 10 min, with 1 μM PKCε agonist (ψεRACK) for 10 min [note 1 in (*5*)], or with 1 μM PKCε antagonist (εV1-2) for 5 min, followed by 10 min of perfusion together with 50 mM ethanol. The hearts were then subjected to a 30-min period of no-flow ischemia before homogenization. Treatment with ethanol and ψεRACK induced a leftward shift of ALDH2 as compared with control, which was blocked with εV1-2 treatment. Blots were probed with antibodies against ALDH2 or against phospho-Ser and phospho-Thr (*5*). **(B)** ALDH2 activity correlates with cardiac protection from ischemic injury. Measurements of ALDH activities in normoxic and ischemic rat hearts treated with ethanol (EtOH, 50 mM), PKCε agonist (ψεRACK), or PKCε antagonist (εV1-2) in the presence of ethanol using the Langendorf apparatus (*5*). Ischemic hearts were also treated with the ALDH2 inhibitor cyanamide (CYA) in the presence or absence of ethanol, PKCε agonist and antagonist, and the ALDH2 desensitizer, nitroglycerin (GTN). Shown is ALDH2 activity (μmol of NADH/min per mg protein) as a function of infarct size, measured by 2,3,5-triphenyltetrazolium chloride (TTC) staining from corresponding heart samples derived from the same studies as in Table 1. Linear regression yielded a high inverse correlation of $R^2 = 0.95$.

[1]Department of Chemical and Systems Biology, Stanford University School of Medicine, Stanford, CA 94305–5174, USA. [2]Department of Biochemistry and Molecular Biology, Indiana University School of Medicine, Indianapolis, IN 46202, USA.

*To whom correspondence should be addressed. E-mail: mochly@stanford.edu

the mitochondria and phosphorylate ALDH2 directly.

We next determined whether ALDH2 is activated in the intact heart following PKCε activation or ethanol treatment and whether there is a correlation between the activity of ALDH2 and infarct size under various treatment conditions. Ischemia alone did not affect ALDH2 activity (Table 1). However, ethanol treatment caused a 21% increase in ALDH2 activity relative to control and a 27% reduction in infarct size [$P < 0.05$ (Table 1 and Fig. 1B)]. Treatment with the selective PKCε activator ψεRACK (*8*) increased ALDH2 activity by 33% with a concomitant 49% reduction in infarct size; and inhibition of PKCε by the selective antagonist εV1-2 (*7*) abolished both the ethanol-induced increase in ALDH2 activity and the ethanol-induced cardiac protection from ischemia (Table 1). Further, in the presence of the ALDH inhibitor cyanamide (5 mM) (*5*, *10*), ALDH2 activity was inhibited by 63% and infarct size increased by 49%, without causing cardiac damage under normoxic conditions; cyanamide also abolished ethanol- or ψεRACK-induced protection and ALDH2 activation (Table 1 and Fig. 1B).

Because cyanamide inhibits several ALDHs, we used another means to inhibit ALDH2. ALDH2 metabolizes nitroglycerin, which leads to generation of the vasodilator, nitric oxide. Yet prolonged treatment with nitroglycerin decreases ALDH2 activity (*5*, *11*). We reasoned that if ALDH2 activity is critical for cardioprotection from ischemic damage, prolonged treatment with nitroglycerin should inhibit PKCε-dependent preconditioning. As expected, a 30-min treatment of nitroglycerin (GTN, 2 μM) in the ex vivo myocardial infarction model in rodents greatly inhibited ALDH2 activity and abolished ethanol- and PKCε-induced activation of ALDH2 (Table 1 and Fig. 1B), whereas the activity of another cardiac dehydrogenase remained unchanged (fig. S4, A and B) (*5*), which indicated that the changes in ALDH2 activity are probably specific. Concomitantly, GTN treatment increased ischemic cardiac damage from 45% in control to 59% and to 63 or 61% in the presence of ethanol or the PKCε activator (Table 1 and Fig. 1B). This effect was not due to nitric oxide generation; treatment with another nitric oxide–generating vasodilator, sodium nitroprusside (SNP, 10 μM), did not affect ALDH2 activity nor did it result in an increase in infarct size (Table 1). There-

Table 1. ALDH2 activity and infarct size in rat hearts subjected to ischemia and reperfusion, ex vivo. The experimental details are provided in (*5*) and in Fig. 1. EtOH, ethanol; ψεR, ψεRACK.

Treatment	ALDH2 activity (μmol NADH/min per mg)	Infarct size (%)	*n*
Normoxia	2.5 ± 0.1	5 ± 4	3
+ CYA	0.9 ± 0.1*	7 ± 5	3
Ischemia	2.4 ± 0.2	45 ± 6	5
+ EtOH	2.9 ± 0.1**	33 ± 10**	5
+ ψεR	3.2 ± 0.1**	23 ± 10†	5
+ εV1-2	2.4 ± 0.3	42 ± 5	5
+ CYA	0.9 ± 0.1†	67 ± 9†	5
+ CYA + EtOH	0.9 ± 0.1†	73 ± 8†	5
+ CYA + ψεR	0.6 ± 0.1†	70 ± 7†	5
+ GTN	1.7 ± 0.1†	59 ± 8**	8
+ GTN + EtOH	1.5 ± 0.1†	63 ± 9**	8
+ GTN + ψεR	1.5 ± 0.1†	61 ± 6**	8
+ SNP	2.3 ± 0.1	45 ± 7	8
+ GTN off	2.1 ± 0.1	33 ± 8**	5
+ GTN on	1.5 ± 0.1**	59 ± 7**	5

*$P < 0.01$ from normoxia. **$P < 0.05$ from ischemia. †$P < 0.01$ from ischemia.

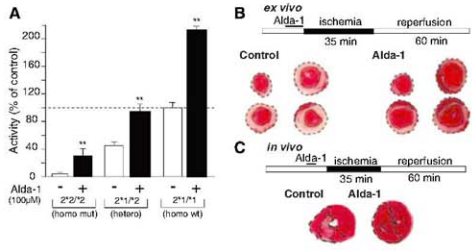


Fig. 2. (A) Alda-1 increases ALDH2 activity. Activation of wild-type and mutant versions of ALDH2 (homotetramers and heterotetramers) by Alda-1 (100 μM). Enzymatic activity of recombinant ALDH2 proteins (20 μg each) is presented as a percentage of control [$n = 3$; **$P < 0.01$ versus control; (*5*)]. **(B)** Alda-1 reduces cardiac damage in an ex vivo model of ischemia and reperfusion injury. (Top) Ex vivo cardiac ischemia model protocol. Myocardial infarct size, induced by 35 min of ischemia followed by 60 min of reperfusion after 10-min pretreatment with Alda-1 (20 μM) or vehicle control using Langendorff apparatus, as in Fig. 1, A and B [$n = 6$; $P < 0.05$; (*5*)]. Representative cross-sectional slices derived from a single heart stained by TTC without (control) and with Alda-1 treatment. Infarct area is indicated by the light pink color and marked with dotted lines. **(C)** Alda-1 reduces cardiac damage in an in vivo rat model of acute myocardial infarction. (Top) In vivo cardiac ischemia model protocol. Reduction of infarct size by injection of Alda-1 (8.5 mg/kg) before left anterior descending coronary artery (LAD) ligation (*5*) was also determined in vivo [$n = 7$; $P < 0.01$ (fig. S6, A and B)]. Shown is TTC staining of representative cross-sectional slices (seven rats per group).

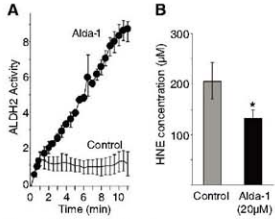


Fig. 3. (A) Effect of Alda-1 on 4HNE metabolism by ALDH2. In vitro metabolism of 4HNE (200 μM) by ALDH2 (arbitrary units) is lost within 1 min of incubation with the substrate (*5*), presumably because of 4HNE-induced ALDH2 inactivation (*18*). 4HNE-induced ALDH2 inactivation is blocked by Alda-1 (20 μM) ($n = 3$) as compared with vehicle control ($n = 3$). **(B)** The protection of ALDH2 from 4HNE-induced inactivation by Alda-1 correlates with a 34% reduction in 4HNE levels ($n = 4$; $P < 0.05$).

fore, there is an inverse correlation between ALDH2 activity and cardiac damage [$R^2 = 0.95$ (Fig. 1B)], strongly suggesting that ALDH2 plays a pivotal positive role in mediating cardiac protection against ischemic injury. Creatine phosphokinase (CPK) release from the heart as an indicator of cardiac damage (*4*) yielded similar results [$R^2 = 0.97$ (fig. S4, C and D)].

Nitroglycerin confers cardiac protection if the prolonged nitroglycerin treatment is terminated at least 1 hour before the ischemic event (*12*). Consistent with these findings, we found that 13 hours of nitroglycerin treatment (5 μg/min per kg of body weight, delivered by a patch) that was terminated 3 hours before the ischemic event decreased cardiac infarct size from 45 to 33% [GTN-off (Table 1)]. However, similar to our ex vivo data, when the nitroglycerin patch was left on, infarct size increased from 45 to 59% [GTN-on (Table 1)]. Therefore, sustained nitroglycerin treatment increased ischemic damage, probably by inducing ALDH2 inactivation [note 4 in (*5*)].

The inverse correlation between ALDH2 activity and cardioprotection against ischemic damage in rat [$R^2 = 0.95$ (Fig. 1B)] (fig. S4D) does not prove that ALDH2 activation is sufficient to induce cardioprotection. We therefore searched for ALDH2 agonists using a high-throughput screen [note 1 in (*5*)] and identified *N*-(1,3-benzodioxol-5-ylmethyl)-2,6-dichlorobenzamide (Alda-1, $M_r = 324$) and similar analogs as ALDH2 activators (Fig. 2A and fig. S5A). We next determined whether Alda-1 activates ALDH2*2, a common East Asian mutant form that has only 1 to 5% of the catalytic activity of the wild-type ALDH2*1 form. [This E487K mutation (Glu at position 487 is replaced by Lys) is at the interface of the tetramer (*13*).] Alda-1 [median effective concentration (EC_{50}) ≈ 20 μM] increased the activity of the mutant, ALDH2*2, 11-fold, the heterotetramer 2.2-fold (similar to the base levels of wild-type ALDH2), and the wild-type ALDH2*1/*1 homotetramers 2.1-fold over basal activity (Fig. 2A and fig. S5, C and D). Alda-1 had no effect on the activity of alcohol dehydrogenase 1 (*5*, *14*), the cytosolic aldehyde dehydrogenase, ALDH1 (*15*), or the mitochondrial enzyme ALDH5 (*5*, *16*) (fig. S5B [note 5 in (*5*)]).

We next used Alda-1 to determine whether direct ALDH2 activation was sufficient to induce cardioprotection. Rat hearts treated ex vivo with 20 μM Alda-1 before 35 min of ischemia followed by 60 min of reperfusion (as in Fig. 2B) had a 26 ± 6% smaller infarct (Fig. 2B) and 24 ± 7% less CPK release ($n = 6$; $P < 0.05$). Alda-1 also reduced infarct size in an in vivo rat model of acute myocardial infarction. After 35 min of ischemia and 60 min of reperfusion, infarct size of the left ventricular free wall was 43 ± 4% ($n = 7$) (*5*). Administration of 8.5 mg/kg Alda-1 into the left ventricle 5 min before ischemia decreased the myocardial infarction by 60 ± 4% [$n = 7$, $P < 0.01$ (Fig. 2C)] (fig. S6, A and B). Although low levels of noxious stimuli trigger cardioprotection (*1*, *2*), Alda-1–induced cardioprotection was not associated with such a stress. JNK, a sensitive marker of cell stress, was not activated by Alda-1 treatment [fig. S7 and note 6 in (*5*)]. Therefore, activation of ALDH2 is sufficient to protect the heart from ischemia damage, in vivo.

4-Hydroxynonenal (4HNE) is a toxic aldehyde that accumulates during cardiac ischemia (*17*); thus, its removal by ALDH2 may be, at least in part, the mechanism by which ALDH2 activation protects the heart from ischemic damage. Furthermore, 4HNE itself can inactivate ALDH2 by forming protein adducts with the enzyme, which limits 4HNE removal (*18*). We confirmed that 4HNE induced rapid inactivation of ALDH2 in vitro and found that 4HNE-induced inactivation of ALDH2 was blocked by Alda-1 (Fig. 3A), which increased the detoxification of 4HNE (Fig. 3B). The molecular basis for Alda-1–induced ALDH2 protection is under investigation, but it is probably due to prevention of 4HNE adduct formation on ALDH2 (*18*).

Although some of the pharmacological tools we used to regulate ALDH2 are relatively nonspecific, 12 different conditions demonstrate the correlative relation between ALDH2 activity and infarct size [$R^2 = 0.95$ (Fig. 1B)]. Therefore, our data strongly suggest that ALDH2 activity is critical for cardioprotection from ischemia. In addition, ALDH2 contributes to ethanol metabolism, and ethanol was used to activate ALDH2. However, ethanol metabolism is unlikely to play a role in the ALDH2-mediated protection; ALDH2 activation also occurred in the absence of ethanol, as well as when we used the PKCε-selective activator or Alda-1. Finally, the importance of cytotoxic aldehydes, such as 4HNE, to overall ischemic injury has been previously suggested [note 7 in (*5*)] (*17*, *19*, *20*). It is possible that the major benefit of Alda-1 is to prevent the inactivation of cytoprotective ALDH2 by 4HNE, which would ensure continual detoxification of oxidative stress–induced cytotoxic aldehydes.

Our results raise the possibility that pharmacological enhancement of ALDH2 activity may be beneficial for patients subjected to cardiac ischemia (e.g., during coronary bypass surgery). The ability of Alda-1 to partially complement or restore mutant ALDH2*2 activity is noteworthy, as it is rare to find a small molecule that can specifically rescue a mutation in humans. Finally, our data from rodent models suggest that the prolonged use of nitroglycerin in East Asian carriers of *Aldh2*2* who experience an ischemic event may need to be reconsidered and that these patients may benefit even more than carriers of the wild-type enzyme if treated with ALDH2 activators.

References and Notes

1. R. Bolli, *Am. J. Physiol. Heart Circ. Physiol.* **292**, H19 (2007).
2. J. M. Downey, A. M. Davis, M. V. Cohen, *Heart Fail. Rev.* **12**, 181 (2007).
3. A. Nakano, M. V. Cohen, J. M. Downey, *Pharmacol. Ther.* **86**, 263 (2000).
4. C. H. Chen, M. O. Gray, D. Mochly-Rosen, *Proc. Natl. Acad. Sci. U.S.A.* **96**, 12784 (1999).
5. Materials and methods, further information on data, and additional discussion are available as supporting material on *Science* Online.
6. M. Jaburek *et al.*, *Circ. Res.* **99**, 878 (2006).
7. M. O. Gray, J. S. Karliner, D. Mochly-Rosen, *J. Biol. Chem.* **272**, 30945 (1997).
8. G. W. Dorn 2nd *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **96**, 12798 (1999).
9. C. Chen, D. Mochly-Rosen, *J. Mol. Cell. Cardiol.* **33**, 581 (2001).
10. R. A. Deitrich, P. A. Troxell, W. S. Worth, *Biochem. Pharmacol.* **25**, 2733 (1976).
11. Z. Chen *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 12159 (2005).
12. M. Hill *et al.*, *Circulation* **104**, 694 (2001).
13. H. N. Larson, H. Weiner, T. D. Hurley, *J. Biol. Chem.* **280**, 30550 (2005).
14. G. Duester *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **81**, 4055 (1984).
15. L. C. Hsu, K. Tani, T. Fujiyoshi, K. Kurachi, A. Yoshida, *Proc. Natl. Acad. Sci. U.S.A.* **82**, 3771 (1985).
16. P. Blasi *et al.*, *Mol. Genet. Metab.* **76**, 348 (2002).
17. P. Eaton, J. M. Li, D. J. Hearse, M. J. Shattock, *Am. J. Physiol.* **276**, H935 (1999).
18. J. A. Doorn, T. D. Hurley, D. R. Petersen, *Chem. Res. Toxicol.* **19**, 102 (2006).
19. D. T. Lucas, L. I. Szweda, *Proc. Natl. Acad. Sci. U.S.A.* **95**, 510 (1998).
20. J. Chen, G. I. Henderson, G. L. Freeman, *J. Mol. Cell. Cardiol.* **33**, 1919 (2001).

21. We thank K. Inagaki and C. L. Murriel for initial studies; D. Solow-Cordero for the high-throughput screen; and F. Conti, D. Cheng, and A. Masoud Sadaghiani for the biochemical characterization of Alda-1. Supported by NIH grant AA11147 (to D.M.-R.) and by Stanford's SPARK Program (C.-H.C.). D.M.-R. is the founder and a shareholder of KAI Pharmaceuticals; however, none of the work described here was done in collaboration with, or with support from, this company. C.-H.C., T.D.H., and D.M.-R. are listed as inventors in three pending patents filed by Stanford University that are related to this work.

Supporting Online Material

www.sciencemag.org/cgi/content/full/321/5895/1493/DC1
Materials and Methods
SOM Text
Figs. S1 to S7
References

1 April 2008; accepted 12 August 2008
10.1126/science.1158554

Dual Origin of Tissue-Specific Progenitor Cells in *Drosophila* Tracheal Remodeling

Molly Weaver and Mark A. Krasnow*

During *Drosophila* metamorphosis, most larval cells die. Pupal and adult tissues form from imaginal cells, tissue-specific progenitors allocated in embryogenesis that remain quiescent during embryonic and larval life. Clonal analysis and fate mapping of single, identified cells show that tracheal system remodeling at metamorphosis involves a classical imaginal cell population and a population of differentiated, functional larval tracheal cells that reenter the cell cycle and regain developmental potency. In late larvae, both populations are activated and proliferate, spread over and replace old branches, and diversify into various stalk and coiled tracheolar cells under control of fibroblast growth factor signaling. Thus, *Drosophila* pupal/adult tissue progenitors can arise both by early allocation of multipotent cells and late return of differentiated cells to a multipotent state, even within a single tissue.

Drosophila larval tissues are composed of differentiated larval cells and imaginal cells. Imaginal cells are pupal and adult tissue progenitors that reside in clusters embedded in or attached to larval tissue (*1*). They remain quiescent during embryogenesis and part or all of larval life, then proliferate and differentiate into pupal and adult tissues at metamorphosis (*2*). By contrast, larval cells cease dividing and differentiate early in development; however, they typically enlarge and become polyploid during larval life (*3*). At metamorphosis, most larval cells die (*4*). Although some larval neurons (*5*) and muscles (*6*) are retained in adult tissues, no differentiated cells are known to reenter the cell cycle and generate new cells and tissues. We show that tracheal (respiratory) system remodeling at metamorphosis is carried out by a classical imaginal cell population and another progenitor population that, like facultative stem cells in mammals (*7*), arises from differentiated cells.

During embryogenesis, the tracheal system develops from segmentally repeated groups of ~80 cells that express Trachealess transcription factor and invaginate, forming sacs attached to epidermis by a stalk of spiracular branch (SB) cells (*8–10*). Branches bud from the sacs and cells diversify primarily under control of Branchless FGF (fibroblast growth factor), which activates Breathless FGFR (FGF Receptor) on tracheal cells (*11, 12*). At metamorphosis, posterior tracheal segments Tr6 to Tr10 are lost (Fig. 1A); new branches form in Tr4 and Tr5 to supply posterior tissues and in Tr2 to supply flight muscle (*8, 13–16*). Although most branches in Tr1 to Tr5 are retained, most of their cells are replaced by imaginal cells (*17, 18*).

Howard Hughes Medical Institute and Department of Biochemistry, Stanford University School of Medicine, Stanford, CA 94305–5307, USA.

*To whom correspondence should be addressed. E-mail: krasnow@cmgm.stanford.edu

Previous work indicated that imaginal tracheal cells (tracheoblasts) compose the SB (Fig. 1B). Unlike other tracheal cells, SB cells express imaginal marker *escargot* (*esg*), remain small and quiescent during embryonic and early larval life, and do not form gas transport tubes (Fig. 1, C and G) (*8, 15, 17, 19–21*). Bromodeoxyuridine (BrdU) incorporation studies showed that SB cells enter S phase at the beginning of the third larval period (L3) and divide 12 to 16 hours later (*17*). Proliferation continues for 24 hours, generating an expanding cluster of tracheoblasts at the SB-transverse connective (TC) junction (Fig. 1, D and E). SB cells in the embryo and early larva express *trachealess* (*trh*) but, unlike most other tracheal cells (*22*), do not express the Trachealess target gene *breathless* (*btl*) (Fig. 1C): The tracheal program is apparently arrested at this step. When activated in L3, they turn down *esg* and turn on *btl* as they proliferate and leave the SB (Fig. 1, D to F).

We developed a "molecular timer" strain (*23*) that highlights the burst of *btl* expression in activated tracheoblasts, which allowed us to distinguish them from the larval tracheal cells they migrate over and replace (Fig. 1, D to F). SB tracheoblasts followed stereotyped paths. In Tr4, they migrated along the TC onto the visceral branch (Fig. 1F'); later, some differentiated into coiled tracheolar (CT) cells (see Fig. 1J). Tracheoblasts respected specific boundaries, never spreading into neighboring tracheal segments or populating the dorsal trunk (DT) (Fig. 1F'). However, tracheoblasts were observed on the other side of the DT, along the dorsal branch (DB) of Tr4 and other anterior DBs (Fig. 1F') (*17*).

If DB tracheoblasts arise from SB tracheoblasts, they would have to move across the DT to reach the DB. However, tracheoblasts were never seen crossing the DT. To exclude this possibility, the fate of SB tracheoblasts was mapped using SB-specific flipase (FLP) recombinase to permanently label SB cells and their descendants (Fig. 1, G to J). This labeled all tracheoblasts migrating out of the SB (Fig. 1, H and J), but not DB tracheoblasts (Fig. 1, H and I). Hence, DB tracheoblasts arise independently.

To identify the source of DB tracheoblasts, we scrutinized early L3 larval DBs, which comprise five to seven cells (Fig. 1B) (*21*), but found no additional cells or cells with the distinctive small size and nuclear morphology of SB tracheoblasts. The positions and number of tracheoblast clones in a clonal analysis of larval DBs (fig. S1 and SOM Text) suggested that DB tracheoblasts arise from ~4 to 5 progenitors along the DB.

We considered whether differentiated stalk cells (DB3 to DB7) might be the source of DB tracheoblasts (*15, 17*). To test this, we used a heat-inducible FLP transgene to permanently label and trace the fate of individual tracheal cells identified in live L2 larvae (Fig. 2, A to C, Table 1, and table S1). This demonstrated that larval DB stalk cells are the source. Individual stalk cells displayed a range of proliferative capacities, giving rise to 2 to 22 tracheoblasts (8.3 ± 7.0, mean ± SD), although occasionally a labeled stalk cell failed to proliferate or degenerated, the standard fate of DB1 and DB2 cells (Table 1 and fig. S2).

DB stalk cells have a complex morphology unlike typical progenitor or stem cells: They are tubular, with autocellular junctions, some (DB3 cells) forming Y-shaped tubes (Fig. 1B). Yet these cells become proliferating, migrating DB tracheoblasts while maintaining contacts with neighboring tracheal cells (Fig. 2, D to F).

Phosphohistone H3 staining (Fig. 2, D to F) showed that DB stalk cells in anterior segments begin dividing 14 to 16 hours after the second molt. Even before they divide (Fig. 2G), they have smaller nuclei than other larval tracheal cells, including DB stalk cells in posterior segments (Fig. 2H), which are otherwise indistinguishable but do not give rise to tracheoblasts. BrdU labeling of newly molted third-instar larvae showed that anterior DB stalk cells do not incorporate the label (Fig. 2G), implying that they do not endoreplicate and presumably remain diploid, unlike posterior DB cells and other differentiated larval cells (Fig. 2, H to J), most of which endoreplicate and become polyploid (*3*).

Although most DB tracheoblasts form multicellular stalks of pupal DBs (Table 1), in Tr2 they form more elaborate structures (Fig. 3, A to D). After proliferating and spreading along the DB (Fig. 3, A and B), they aggregate (Fig. 3B), form secondary branches (Fig. 3C), and differentiate into multicellular stalks (MS), unicellular stalks (US), and Blistered (DSRF)–expressing CT cells with coiled intracellular lumens (Fig. 3, C and D, and fig. S3) that unfurl on flight muscle (*8, 24*). Fate mapping showed that single DB stalk cells in Tr2 routinely formed mixed clones containing MS, US, and CT cells (Fig. 3, E to G, Table 1, and table S1). Thus, DB stalk cells in Tr2 transform into multipotent tracheoblasts that can proliferate and acquire different fates.

Bnl/Btl signaling controls cell fate selection in the embryo (*11, 12*). To test for function in pupal tracheoblasts, we generated *btl* clones. These rare-

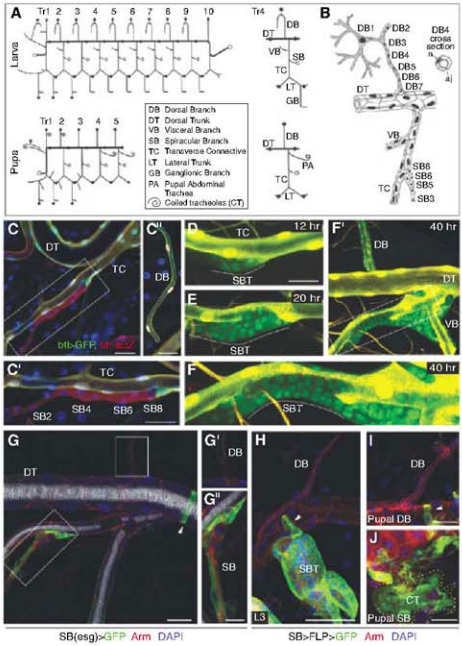


Fig. 1. Spiracular branch tracheoblasts are not the source of dorsal branch tracheoblasts. **(A)** *Drosophila* tracheal remodeling at metamorphosis. (Left) Schematics of larval and pupal tracheal system. Tr, tracheal metamere. Filled circles, positions of fusion cells connecting metameres. Open circles, attachments to epidermis. (Right) Branch names in Tr4. **(B)** Cellular structure of larval Tr4. DB contains terminal (DB1), fusion (DB2), and several stalk (DB3 to 7) cells. SB contains six to eight cells (SB1 to 8). SBs are collapsed but open at molts for tracheal cuticle extrusion. n, nucleus; aj, autocellular junction. **(C)** SB in early L3 larva carrying *trh-lacZ* (antibody to beta-galactosidase, red), *btl>GFP* (*btl-Gal4;UAS-GFP*) [antibody to green fluorescent protein (GFP), green]. Nuclei are stained with 4′,6′-diamidino-2-phenylindole (DAPI) (blue). (**C'** and **C''**) SB (boxed) and DB in same segment. All tracheal cells, including SB and DB cells, express *trh-lacZ*; all except SB cells express *btl>GFP*. (**D** to **F**) SB in "molecular timer" strain (*btl-Gal4,UAS-GFP;UAS-DsRed*) at indicated times in L3. (D) *btl-Gal4* expression initiating in SB tracheoblasts (SBTs) visualized by upstream activating sequence (UAS)–GFP expression (green). DsRed (red) takes longer to mature, so SBTs appear green, whereas tracheal cells in which *btl-Gal4* was previously active express both proteins, so appear yellow. (E and F) Proliferating and spreading SBTs. (**F'**) Low magnification of (F). Figure S5 shows low magnification of (D) and (E). **(G)** Early L3 *esg>GFP* (*esg-Gal4,UAS-GFP*) larva immunostained for GFP (green) and Armadillo (cell junction marker, red). Blue, DAPI-stained nuclei. White (reflected light), air-filled lumen. Insets (**G'** and **G''**), DB and SB (boxed). SB cells, but not DB or other tracheal cells except fusion cells (arrowhead), express *esg*. (**H** to **J**) Lineage tracing of SB (*esg*-expressing) cells with *P127-Gal4;UAS-FLP;yAc-GFP*. SBTs and fusion cells (arrowheads) express GFP lineage marker in late L3 (H) and pupae (I and J) where some SB tracheoblasts (e.g., PA) form CTs [outlined in (J)]. DB tracheoblasts do not express lineage marker (H and I). Bars, 25 μm, except (G') and (G''), 10 μm.

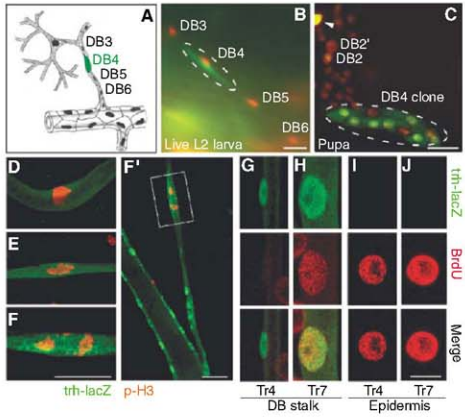


Fig. 2. DB stalk cells proliferate and differentiate into tracheoblasts. **(A** to **C)** Fate mapping a DB4 stalk cell (clone O, Table 1). (A) Schematic showing labeled cell. (B) DB4 cell labeled with cytoplasmic GFP. All tracheal nuclei express DsRed (*btl-Gal4;UAS-nuclearDsRed*). (C) Same DB three days later. Labeled DB4 generated seven-cell clone. DB2 and contralateral DB2 (DB2') are indicated along with degenerating contralateral DB1 (arrowhead), which was also labeled in L2 larva. (**D** to **F**) DB4 cells at three different stages of mitosis in L3 larvae immunostained for phosphohistone H3 (red) and beta-galactosidase (*trh-lacZ*, green). Tubular DB4 cell divides along short axis of tube (F and **F'**). Figure S6 shows low magnification of (D) and (E). (**G** to **J**) Individual DB stalk cells in Tr4 and Tr7, and nearby epidermal cells of newly molted *trh-lacZ* L3 larva labeled with BrdU during L1 and L2 then costained after molting for BrdU (red) and beta-galactosidase (green). Figure S7 shows low magnification of (G) to (J). Bars, 25 μm, except (G) to (J), 10 μm.

ly formed CT cells (fig. S3 and table S2). Likewise, conditional expression of dominant-negative Btl in the L3 tracheal system reduced or eliminated secondary branches and CT cells (Fig. 3, H and I). Constitutively active receptor induced ectopic secondary branches and CT cells throughout the tracheal system, including all anterior DBs (Fig. 3J and fig. S4). Thus, DB tracheoblasts in anterior segments can acquire new tracheal fates, and FGF signaling also plays a critical role in reselecting cell fates when DB stalk cells are reactivated.

Anterior DB stalk cells are the first differentiated cells in *Drosophila* shown to reenter the cell cycle and regain developmental potency. They regain the same abilities to proliferate, spread, and differentiate into various tracheal cell types as SB tracheoblasts, classical imaginal cells that remain quiescent—blocked in tracheal outgrowth and cell diversification—during embryonic and most of larval life. This suggests that both types of tracheal progenitors arrive at a similar state, one by early developmental arrest (SB tracheoblasts), the other by late return to an earlier state (anterior DB stalk cells). The only known features that distinguish these cells from tracheal cells that lack progenitor potential (including posterior DB stalk cells that are otherwise indistinguishable from anterior DB stalk cells) are their small nuclear size and lack of endoreplication (table S3). These features may be part of a program that maintains, or allows cells to regain, the proliferative and diversification potential of early tracheal cells. This program is operative in imaginal tracheal cells and can apparently be implemented in other tracheal cells, independent of their differentiation program.

The blurring of the distinction between imaginal and differentiated larval cells in *Drosophila* parallels a current debate about adult stem cells in mammals (*25*). Some mammalian tissues have

Table 1. Fate mapping of single larval dorsal branch cells.

Clone	Marked cell		No. cells in clone	Clone cell fates					Class*
				US	MS	CT	Dead	Lost	
A	DB1	Tr5	1†	0	†	†	1	0	0
E	DB1	Tr2	0	0	0	0	0	1	0
G	DB2	Tr2	0	0	0	0	0	1	0
H	DB3	Tr3	1	0	0	0	1	0	0
I	DB3	Tr2	2	2	0	0	0	0	I
J	DB3	Tr2	5	5	0	0	0	0	I
K	DB3	Tr2	8	3	1	4	0	0	II
L	DB3	Tr2	12	7	1	4	0	0	II
M	DB3	Tr2	21	18	2	1	0	0	II
N	DB4	Tr3	1	0	0	0	1	0	0
O	DB4	Tr2	7	7	0	0	0	0	I
P	DB5	Tr2	0	0	0	0	0	1	0
Q	DB5	Tr3	1	0	0	0	1	0	0
R	DB5	Tr3	14†	14	†	†	0	0	I
S	DB5	Tr2	14	6	2	6	0	0	II
T	DB6	Tr5	1†	0	†	†	1	0	0
U	DB6	Tr3	>6	>6	0	0	0	0	I
V	DB6	Tr2	10	4	0	6	0	0	II
W	DB7	Tr3	14	14	0	0	0	0	I

*Class 0, no proliferation, followed by necrosis or death; Class I, proliferation generating cells of the same fate; Class II, proliferation generating cells of multiple fates. †Clone analyzed at W3L, before cell differentiation.

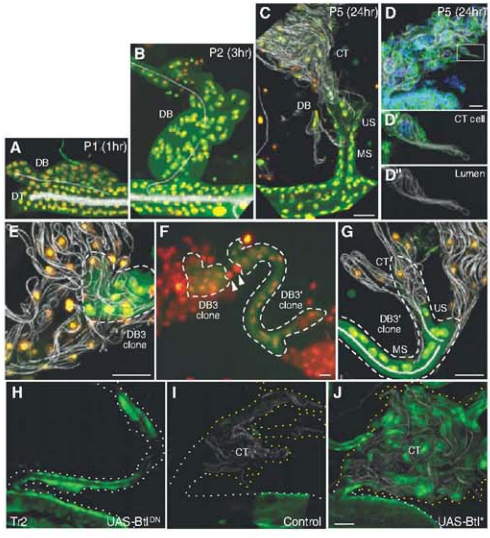


Fig. 3. DB stalk cells are multipotent progenitors. **(A to C)** Tr2 DB tracheoblast morphogenesis in *btl-Gal4;UAS-nDsRed;UAS-GFP* pupae at pupal stages and times after pupariation indicated. Tracheal cytoplasm appears green (GFP), nuclei yellow (GFP, nDsRed), and air-filled lumen white (reflected light). **(D)** CT cell cluster along Tr2 DB of *btl-Gal4/+;UAS-Apc2-GFP/+* pupa showing microtubules (APC-GFP, green), nuclei (DAPI, blue), and air-filled lumen (reflected light, white). **(D' and D")** Boxed CT cell with coiled lumen. **(E to G)** Fate mapping of L2 larva DB3 cell in Tr2 and contralateral DB3 (DB3') analyzed at stage P5 (F). DB3 generated 12-cell clone [(E), clone L in Table 1]; DB3' generated 21-cell clone [(G), clone M in Table 1]. Both clones contain MS, US, and CT cells. **(H to J)** Tr2 tracheal segment of *btl-Gal4,UAS-GFP/tubulin Gal80TS;btl-Gal4,UAS-GFP/UAS-DNbtl* (Btl^{DN}) or *UAS-λbtl* (Btl*) animals grown at 18°C and shifted to 30°C in late L2 to express dominant-negative (H) or constitutively active Breathless (Btl*) (J), or left at 18°C as control (I). Dotted lines, DBs and DTs; yellow dots, regions with CTs. Tracheal cells in experimental animals express GFP; expression in control (I) is inhibited by $Gal80^{TS}$, so tracheal cells were visualized by reflection (white) and DAPI staining. Bars, 25 μm.

dedicated stem cells maintained in a primitive state (*26*). However, other tissues may rely on facultative stem cells, differentiated cells that reenter the cell cycle to replenish lost cells (*7*). Our results show that progenitors with each of these features are present in a single *Drosophila* tissue and that both play crucial roles. This provides a tractable system for dissection of the arrest of a tissue-specific developmental program and reversal to an earlier, more plastic state, important steps in tissue engineering, repair, and cancer.

Note added in proof: Guha *et al.* (*27*) show that differentiated tracheal cells in Tr2 can reenter the cell cycle.

References and Notes

1. R. Harbecke *et al.*, *Int. J. Dev. Biol.* **40**, 197 (1996).
2. P. Kylsten, R. Saint, *Dev. Biol.* **192**, 509 (1997).
3. B. A. Edgar, T. L. Orr-Weaver, *Cell* **105**, 297 (2001).
4. V. P. Yin, C. S. Thummel, *Semin. Cell Dev. Biol.* **16**, 237 (2005).
5. D. W. Williams, J. W. Truman, *J. Neurobiol.* **64**, 24 (2005).
6. R. Klapper, *Mech. Dev.* **95**, 47 (2000).
7. E. L. Rawlins, B. L. M. Hogan, *Development* **133**, 2455 (2006).
8. G. Manning, M. A. Krasnow, in *The Development of Drosophila melanogaster*, A. Martinez, M. Bate, Eds. (CSHL Press, Cold Spring Harbor, NY, 1993), vol. I, pp. 609–685.
9. R. Wilk, I. Weizman, B. Z. Shilo, *Genes Dev.* **10**, 93 (1996).
10. D. D. Isaac, D. J. Andrew, *Genes Dev.* **10**, 103 (1996).
11. E. Zelzer, B. Z. Shilo, *Bioessays* **22**, 219 (2000).
12. A. Ghabrial, S. Luschnig, M. M. Metzstein, M. A. Krasnow, *Annu. Rev. Cell Dev. Biol.* **19**, 623 (2003).
13. J. C. Pihan, *Arch. Zool. Exp. Gen.* **109**, 287 (1968).
14. J. Whitten, *The Genetics and Biology of Drosophila*, M. Ashburner, T. R. F. Wright, Eds. (Academic Press, New York, 1980), vol. 2, pp. 499–540.
15. M. Sato, T. B. Kornberg, *Dev. Cell* **3**, 195 (2002).
16. C. Cabernard, M. Affolter, *Dev. Cell* **9**, 831 (2005).
17. A. Guha, T. B. Kornberg, *Dev. Biol.* **287**, 192 (2005).
18. M. Sato, Y. Kitada, T. Tabata, *Dev. Biol.* **318**, 247 (2008).
19. Materials and methods are available as supporting material on *Science* Online.
20. N. Fuse, S. Hirose, S. Hayashi, *Development* **122**, 1059 (1996).
21. C. Samakovlis *et al.*, *Development* **122**, 1395 (1996).
22. T. Ohshiro, K. Saigo, *Development* **124**, 3975 (1997).
23. A. Ghabrial, M. A. Krasnow, *Nature* **441**, 746 (2006).
24. S. A. Shafiq, *Q. J. Microbiol. Sci.* **104**, 135 (1963).
25. Y. Dor, D. A. Melton, *Cell Cycle* **3**, 1104 (2004).
26. L. L. Weissman, D. J. Anderson, F. Gage, *Annu. Rev. Cell Dev. Biol.* **17**, 387 (2001).
27. A. Guha, L. Lin, T. B. Kornberg, *Proc. Natl. Acad. Sci. U.S.A.* **105**, 10832 (2008).

28. We thank lab members for advice, reagents, and discussion. M.A.K. is an investigator of the Howard Hughes Medical Institute (HHMI); M.W. was an HHMI fellow of the Life Sciences Research Foundation.

Supporting Online Material

www.sciencemag.org/cgi/content/full/1158712/DC1
Materials and Methods
SOM Text
Figs. S1 to S7
Tables S1 to S3
References

4 April 2008; accepted 7 July 2008
Published online 31 July 2008;
10.1126/science.1158712
Include this information when citing this paper.

FBXW7 Targets mTOR for Degradation and Cooperates with PTEN in Tumor Suppression

Jian-Hua Mao,[1]* Il-Jin Kim,[1]* Di Wu,[1] Joan Climent,[1] Hio Chung Kang,[1] Reyno DelRosario,[1] Allan Balmain[1,2]†

The enzyme mTOR (mammalian target of rapamycin) is a major target for therapeutic intervention to treat many human diseases, including cancer, but very little is known about the processes that control levels of mTOR protein. Here, we show that mTOR is targeted for ubiquitination and consequent degradation by binding to the tumor suppressor protein FBXW7. Human breast cancer cell lines and primary tumors showed a reciprocal relation between loss of FBXW7 and deletion or mutation of PTEN (phosphatase and tensin homolog), which also activates mTOR. Tumor cell lines harboring deletions or mutations in *FBXW7* are particularly sensitive to rapamycin treatment, which suggests that loss of *FBXW7* may be a biomarker for human cancers susceptible to treatment with inhibitors of the mTOR pathway.

The *FBXW7* gene (F-box and WD repeat domain–containing 7, also known as hCDC4, FBW7, and hAGO) is a p53-dependent tumor suppressor gene that undergoes deletion and/or mutation in a variety of human tumors (*1*–*3*). Loss or mutation of *FBXW7* has been associated with increased genetic instability or growth deregulation, because of its effects on ubiquitination and turnover of several oncoproteins (*1*–*7*), but the exact mechanism of tumor suppression remains unclear. We carried out a genomewide search for Fbxw7 ubiquitination targets using the consensus CDC4 (cell division cycle 4) phosphodegron (CPD) sequence I/L-I/L/P-T-P-XXXX (where lysine and arginine are unfavorable in the X locations) (*1*, *8*) in a mouse protein database (http://www.ensembl.org/Mus_musculus). This search revealed a strong match within the HEAT domain of mTOR (Frap1) (fig. S1), which regulates cell growth, metabolism, and proliferation. The mTOR CPD region is evolutionarily conserved in different species ranging from humans to zebrafish (fig. S2), which suggests its potential functional importance.

We tested the possibility that Fbxw7 may directly regulate levels of mTOR by analyzing two independent preparations of mouse embryonic fibroblasts (MEFs) from *Fbxw7*$^{-/-}$ mice (Fig. 1A). Depletion of *Fbxw7* increased the levels of both total mTOR and phosphorylated mTOR (p-mTOR), as well as the downstream mTOR target S6-kinase (p-S6K). In contrast, there was no appreciable effect on upstream components of the mTOR signaling pathway such as Akt and phosphorylated Akt (p-Akt) (Fig. 1A).

Depletion of *Fbxw7* in vivo also led to an increase in both mTOR and phosphorylated mTOR (p-mTOR) protein levels. Thymus and spleen from *Fbxw7*$^{+/-}$ mice showed increased mTOR and p-mTOR in comparison with control wild-type tissue (Fig. 1B). In contrast, levels of total Akt and p-Akt (at Ser^{437}) did not change appreciably (Fig. 1B). However, in *Pten*$^{+/-}$ mice (lacking the phosphatase and tensin homolog), both thymus and spleen showed an increase in p-mTOR without any obvious change in total mTOR protein level (Fig. 1B). In agreement with these observations, p-Akt levels were elevated in thymus and spleen from *Pten*$^{+/-}$ mice and, presumably, were responsible for activation of mTOR signaling (Fig. 1B).

A similar trend was seen in human cells from which the *FBXW7* gene had been deleted by homologous recombination (*4*). Both HCT116 and DLD1 cells that lacked *FBXW7* had slightly higher levels of mTOR and p-mTOR, as well as increased levels of p-S6K (fig. S3). Again, no effect was seen on the upstream targets AKT and p-AKT (fig. S3). Human 293T cells transfected with a dominant-negative form of *FBXW7* (HA-FBXW7-AF) (*9*) also showed an increase in the level of both total mTOR and p-mTOR (Fig. 1C), whereas cells expressing the transfected normal hemagglutinin (HA)-tagged FBXW7 (HA-FBXW7) protein reproducibly had lower levels of both total mTOR and p-mTOR (Fig. 1C). In contrast, levels of AKT and p-AKT were not affected by overexpression of *FBXW7* (Fig. 1C).

We next investigated possible interactions between mTOR and FBXW7 proteins. A vector encoding HA-tagged FBXW7 was transfected into 293T cells, followed by immunoprecipitation using antibodies directed against the HA tag or mTOR, and immunoblot analysis. Both showed an interaction between the mTOR and FBXW7 proteins (Fig. 2A). Immunoprecipitation also revealed an interaction between mTOR and FBXW7-ΔF (Fig.

[1]Cancer Research Institute, University of California at San Francisco, 2340 Sutter Street, San Francisco, CA 94143, USA. [2]Department of Biochemistry and Biophysics, University of California at San Francisco, San Francisco, CA 94143, USA.

*These authors contributed equally to this work.
†To whom correspondence should be addressed. E-mail: abalmain@cc.ucsf.edu

2A), indicating that the WD40 domain of *FBXW7* is the interaction site. Additional controls for the specificity of the interaction between FBXW7 and mTOR are shown in figs. S4 and S5.

We cloned a fragment of mTOR (from amino acids 1 to 898) that contained the putative CPD, and generated two mutants, one carrying a deletion of Tyr^{631} [mTOR(delT631)], and the other a point mutation in which Tyr^{631} is replaced by Gly [mTOR(T631G)]. The wild-type and mutant mTOR fragments were coexpressed with HA-FBXW7 in 293T cells, followed by immunoprecipitation with antibodies against one protein and immunoblot analysis of the second. Although the wild-type fragment of mTOR immunoprecipitated efficiently with FBXW7, binding to the mTOR (delT631) and mTOR(T631G) mutants was dramatically reduced (Fig. 2B). Residual binding could be due to the presence of an additional, but weaker, consensus binding site in mTOR (located at position amino acids 314 to 319). Thus, FBXW7 binds mTOR predominantly through the major conserved CPD site.

To determine whether the regulation of mTOR by FBXW7 is through the proteasome-dependent degradation pathway, MCF7 and SUM149PT cells (*10*) (table S1) were treated with the proteasome inhibitor MG-132. Proteasome inhibition caused a dramatic increase in the mTOR levels in MCF7 cells, which retain *FBXW7*, but not in SUM149PT cells, which have a homozygous mutation (Fig. 2C), which suggested that degradation of mTOR is FBXW7-dependent. To examine whether the ubiquitination status of mTOR was FBXW7-dependent, 293T and SUM149PT were transfected with a vector containing CMV promoter-driven HA-ubiquitin. Immunoprecipitation of mTOR followed by immunoblot analysis of the HA-ubiquitin showed that mTOR ubiquitination was found only in cells retaining a functional *FBXW7* gene (fig. S6). The same experiment in HCT116 WT and HCT116 $FBXW7^{-/-}$ cells (Fig. 2D) again showed efficient ubiquitination of mTOR only in HCT116

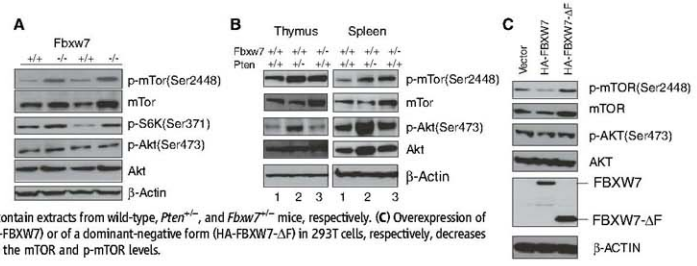


Fig. 1. Depletion of FBXW7 increases mTOR and p-mTOR levels. **(A)** The levels of mTOR and p-mTOR, as well as the downstream mTOR target S6-kinase (P-S6K), are up-regulated in $Fbxw7^{-/-}$ MEFs, whereas the levels of Akt and p-Akt (bottom) were not appreciably affected. Two independent preparations were used. **(B)** Elevation of mTOR in thymus and spleen from $Fbxw7^{+/-}$ mice. Lanes 1 to 3 contain extracts from wild-type, $Pten^{+/-}$, and $Fbxw7^{+/-}$ mice, respectively. **(C)** Overexpression of either full-length FBXW7 (HA-FBXW7) or of a dominant-negative form (HA-FBXW7-ΔF) in 293T cells, respectively, decreases (lane 2) or increases (lane 3) the mTOR and p-mTOR levels.

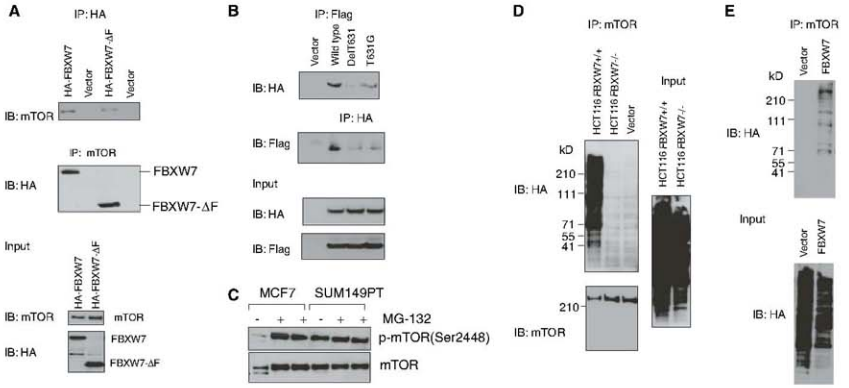


Fig. 2. mTOR interacts with FBXW7. **(A)** Immunoprecipitation (IP) of HA-FBXW7 identifies mTOR as an interacting protein. HA-FBXW7 or FBXW7-ΔF was expressed in human 293T cells, followed by immunoprecipitation of the proteins with HA-specific antibodies and immunoblotting (IB) with antibodies against mTOR (top). The reciprocal experiment shows that the region of *FBXW7* that interacts is the WD40 domain (middle). (Bottom) The input levels of mTOR and FBXW7 proteins in the lysates. **(B)** FBXW7 binds the wild-type fragment of mTOR, but binding to the mTOR(delT631) and mTOR (T631G) mutants was dramatically reduced. **(C)** MCF7 breast cancer cells show increased mTOR levels after treatment with MG-132, but this is not seen in SUM149PT cells, which have no functional *FBXW7* gene. **(D)** HCT116 WT and HCT116 $FBXW7^{-/-}$ cells were transfected with HA-ubiquitin. Immunoprecipitation of mTOR followed by immunoblot analysis of the HA-ubiquitin showed that ubiquitination of mTOR was only seen in the HCT116 WT cells, and not in HCT116 $FBXW7^{-/-}$ cells. The vector lane shows HCT116 $FBXW7^{-/-}$ cells transfected with empty vector construct. **(E)** Ubiquitination of mTOR is restored by exogenous *FBXW7* expression. HCT116 $FBXW7^{-/-}$ cells were transfected with an *FBXW7*-expressing construct and HA-tagged ubiquitin. Immunoprecipitation of mTOR showed increased ubiquitination compared with controls (fig. S7).

Fig. 3. Genetic interaction between *FBXW7* and *PTEN* in human breast cancers. The green bar indicates loss, the red bar indicates gain, and the black bar indicates no changes. **(A)** The data from 53 human breast cancer cell lines, ordered in the vertical axis from 1 to 53 (*11*). The copy number of *FBXW7* and *PTEN* was determined by quantitative PCR with TaqMan. Deletion of *FBXW7* rarely occurred in tumors that also show deletion of *PTEN* ($P = 0.014$). (**B**, **C**, and **D**) Three independent sets of human primary breast cancers, analyzed by the same BAC CGH microarray platform: (B) 185 human primary breast cancers (*12*), (C) 145 human primary breast cancers (*13*), and (D) 67 human primary breast cancers (*14*). The copy number was determined on the basis of published CGH data (*FBXW7* was based on BAC RP11-73G16, *PTEN* on BAC RP11-380G5). Loss is defined as $\log_2(\text{ratio}) < -0.25$ and gain as $\log_2(\text{ratio}) > 0.25$.

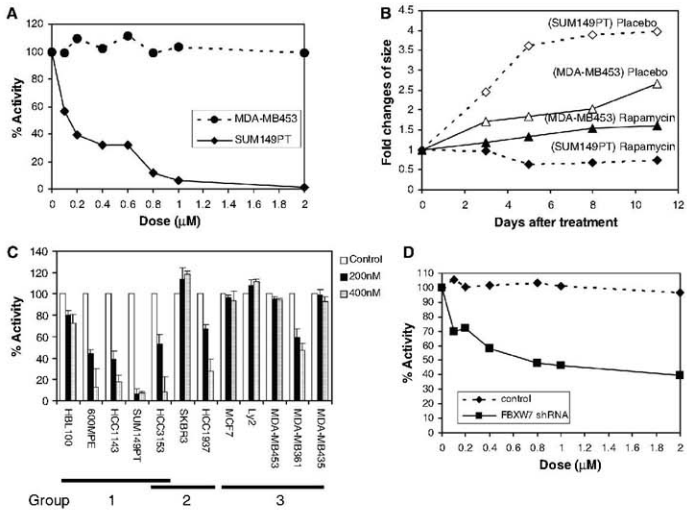


Fig. 4. Loss of FBXW7 increases sensitivity to an mTOR inhibitor (rapamycin). **(A)** The breast cancer SUM149PT cells, which have a homozygous mutation in *FBXW7*, were killed at a rapamycin concentration of 100 nM, whereas MDA-MB453 cells with wild-type *FBXW7* were resistant. **(B)** Treatment of nude mouse xenografts with rapamycin. The SUM149PT cells showed a relative decrease in size, followed by stable tumor growth, whereas the MDA-MB453 cells were unaffected by treatment. **(C)** Sensitivity to mTOR rapamycin in a range of breast cancer cell lines. Tumor cells with deletion or mutation of *FBXW7* are in group 1, those with deletion or mutation of *PTEN* are group 2, and cells with wild-type copies of both genes are in group 3. **(D)** Down-regulation of FBXW7 with specific shRNA in rapamycin-resistant MDA-MB453 cells increases the sensitivity to this treatment.

WT cells. Finally, we cotransfected SUM149PT cells with constructs encoding both FBXW7 and HA-ubiquitin, and found that ubiquitination of mTOR was restored by exogenous *FBXW7* expression (Fig. 2E). Thus, ubiquitination of mTOR is largely, if not exclusively, mediated by binding to FBXW7.

As *FBXW7* and *PTEN* both affect signaling through mTOR, we examined the genetic status of both genes in a panel of 53 breast cancer cell lines (*11*). Quantitative TaqMan real-time polymerase chain reaction (PCR) assays of the number of copies of *FBXW7* and *PTEN* genes in each of the cell lines were in good concordance with data found by bacterial artificial chromosome (BAC) comparative genomic hybridization (CGH) microarray (see table S1). Most of the breast cancer cell lines that exhibited loss of a single copy of *FBXW7* (23 out of 53, Fig. 3A) did not show corresponding loss of *PTEN*. In contrast, of the 14 lines that showed loss of a single copy of *PTEN* (Fig. 3A), only one had also lost a copy of *FBXW7*, which suggested that *FBXW7* and *PTEN* show some functional redundancy in tumor development. Similar results were obtained by examination of the copy number status of genomic regions containing *FBXW7* and *PTEN* genes in three independent human primary breast cancer sets for which BAC CGH microarray data were available (*12*–*14*). From a total of 450 tumor and cell line DNA samples shown in Fig. 3, A to D, only 4 had lost a copies of the regions containing both genes, a result that is unlikely to be a consequence of random genetic alterations ($P = 4.9 \times 10^{-7}$).

We also considered the possibility that other somatic changes such as point mutations or gene-silencing events could affect the results. The *FBXW7* gene continued to be expressed in all 25 breast cancer cell lines examined (fig. S8), which indicated that no gene silencing had occurred, although very low levels were found in five cell lines [lanes 10, 13, 14, 16, and 20 (fig. S8)]. All of these lines had lost one copy of the *FBXW7* gene except one (SUM149PT, lane 16), in which a point mutation was detected (table S1). The *PTEN* gene was found to be silent in two cell lines (fig. S8, lanes 11 and 12), and both had lost one copy of the *PTEN* gene. Three mutations in *PTEN* were found (fig. S8 and table S1). Thus, gene silencing (for example, by promoter methylation) or point mutations in *FBXW7* and *PTEN* are relatively rare mechanisms of inactivation of these genes, in comparison with single-copy deletions. These data are further compatible with the identification of both genes as haplo-insufficient tumor suppressors (*3*, *15*, *16*).

Because deletion or mutation of *FBXW7* in human breast cancer cells leads to increased levels of mTOR, we tested the possibility that cells harboring these deletions may show increased sensitivity to the mTOR inhibitor rapamycin. We treated two breast cancer cell lines, SUM149PT cells (homozygous *FBXW7* mutations) and MDA-MB453 cells (wild-type *FBXW7*) with rapamycin and counted numbers of viable cells. SUM149PT cells proved to be very sensitive to this treatment [median inhibitory concentration (IC_{50}) < 200 nM], whereas MDA-MB453 cells were relatively resistant (IC_{50} > 2 μM) (Fig. 4A). In nude mouse xenografts, groups of five mice were injected with both cell lines, one on each flank, and were treated by intraperitoneal injection with rapamycin over an 11-day period. The SUM149PT cells showed a relative decrease in size followed by stable tumor growth, whereas the MDA-MB453 cells were relatively unaffected by treatment (Fig. 4B).

An additional set of 10 breast cancer cell lines was treated with rapamycin at concentrations of 200 and 400 nM. Cells with deletion or mutation of *FBXW7* (HBL100, 600MPE, SUM149PT, HCC3153, and HCC1143) or *PTEN* (HCC1937 and HCC3153) showed significant sensitivity to killing by rapamycin, although the magnitude of the effect varied (*17*) (Fig. 4C). To establish a direct link between loss of FBXW7 and rapamycin sensitivity, we down-regulated expression levels of FBXW7 using short hairpin RNA (shRNA) (*18*) in the rapamycin-resistant MDA-MB453 cells, which resulted in an increase in sensitivity to this drug [IC_{50}< 0.8 μM (Fig. 4D)].

Our findings implicate FBXW7 in an evolutionarily conserved pathway that controls regulation of mTOR protein levels. Because FBXW7 is a haploinsufficient tumor suppressor that undergoes heterozygous loss in a substantial proportion of human tumors, the data suggest new approaches to reduce mTOR levels in cancers by the use of drugs that may reactivate the remaining copy of FBXW7 in a similar way that nutlins (small-molecule MDM2-antagonists) have been shown to activate wild-type copies of p53 in human tumors (*19*). Loss of FBXW7 may also be a useful biomarker for sensitivity of human tumors to inhibitors of the mTOR pathway.

References and Notes

1. M. Welcker, B. E. Clurman, *Nat. Rev. Cancer* **8**, 83 (2008).
2. S. Akhoondi *et al.*, *Cancer Res.* **67**, 9006 (2007).
3. J. H. Mao *et al.*, *Nature* **432**, 775 (2004).
4. H. Rajagopalan *et al.*, *Nature* **428**, 77 (2004).
5. M. Yada *et al.*, *EMBO J.* **23**, 2116 (2004).
6. M. Welcker *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **101**, 9085 (2004).
7. Z. Kemp *et al.*, *Cancer Res.* **65**, 11361 (2005).
8. Single-letter abbreviations for the amino acid residues are as follows: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; Y, Tyr; and X, any amino acid.
9. G. Wu *et al.*, *Mol. Cell. Biol.* **21**, 7403 (2001).
10. H. Strohmaier *et al.*, *Nature* **413**, 316 (2001).
11. R. M. Neve *et al.*, *Cancer Cell* **10**, 515 (2006).
12. J. Climent *et al.*, *Cancer Res.* **67**, 818 (2007).
13. K. Chin *et al.*, *Cancer Cell* **10**, 529 (2006).
14. J. Fridlyand *et al.*, *BMC Cancer* **6**, 96 (2006).
15. J. H. Mao *et al.*, *Oncogene* **22**, 8379 (2003).
16. A. Di Cristofano, B. Pesce, C. Cordon-Cardo, P. P. Pandolfi, *Nat. Genet.* **19**, 348 (1998).
17. L. S. Steelman *et al.*, *Oncogene* **27**, 4086 (2008).
18. M. Welcker, A. Orian, J. E. Grim, R. N. Eisenman, B. E. Clurman, *Curr. Biol.* **14**, 1852 (2004).
19. J. K. Boolamwini *et al.*, *Curr. Cancer Drug Targets* **5**, 57 (2005).

20. We thank B. Vogelstein for providing us with the HCT116 WT, HCT116 *FBXW7*$^{-/-}$, DLD1 wild-type, and DLD1 *FBXW7*$^{-/-}$ cell lines; K.I. Nakayama for providing *Fbxw7* knockout mice and vectors (HA-FBXW7 and HA-FBXW7ΔF); and O. Tetsu for vector encoding HA-ubiquitin. These studies were supported by NCI grant U01 CA84244 and the U.S. Department of Energy (DE-FG02-03ER63630) to A.B., the University of California at San Francisco Research-Evaluation Allocation Committee (REAC) to J.-H.M. A.B. acknowledges support from the Barbara Bass Bakar Chair of Cancer Genetics.

Supporting Online Material

www.sciencemag.org/cgi/content/full/321/5895/1499/DC1
Materials and Methods
Figs. S1 to S8
Table S1

9 July 2008; accepted 11 August 2008
10.1126/science.1162981

Unsupervised Natural Experience Rapidly Alters Invariant Object Representation in Visual Cortex

Nuo Li and James J. DiCarlo*

Object recognition is challenging because each object produces myriad retinal images. Responses of neurons from the inferior temporal cortex (IT) are selective to different objects, yet tolerant ("invariant") to changes in object position, scale, and pose. How does the brain construct this neuronal tolerance? We report a form of neuronal learning that suggests the underlying solution. Targeted alteration of the natural temporal contiguity of visual experience caused specific changes in IT position tolerance. This unsupervised temporal slowness learning (UTL) was substantial, increased with experience, and was significant in single IT neurons after just 1 hour. Together with previous theoretical work and human object perception experiments, we speculate that UTL may reflect the mechanism by which the visual stream builds and maintains tolerant object representations.

When presented with a visual image, primates can rapidly (<200 ms) recognize objects despite large variations in object position, scale, and pose (*1*, *2*). This ability likely derives from the responses of neurons at high levels of the primate ventral visual stream (*3*–*5*). But how are these powerful "invariant" neuronal object representations built

by the visual system? On the basis of theoretical (*6–11*) and behavioral (*12*, *13*) work, one possibility is that tolerance ("invariance") is learned from the temporal contiguity of object features during natural visual experience, potentially in an unsupervised manner. Specifically, during natural visual experience, objects tend to remain present for seconds or longer, while object motion or viewer motion (e.g., eye movements) tends to cause rapid changes in the retinal image cast by each object over shorter time intervals (hundreds of ms). The ventral visual stream could construct a tolerant object representation by taking advantage of this natural tendency for temporally contiguous retinal images to belong to the same object. If this hypothesis is correct, it might be possible to uncover a neuronal signature of the underlying learning by using targeted alteration of those spatiotemporal statistics (*12*, *13*).

To look for such a signature, we focused on position tolerance. If two objects consistently swapped identity across temporally contiguous changes in retinal position then, after sufficient experience in this "altered" visual world, the visual system might incorrectly associate the neural representations of those objects viewed at different positions into a single object representation (*12*, *13*). We focused on the top level of the primate ventral visual stream, the inferior temporal cortex (IT), where many individual neurons possess position tolerance—they respond preferentially to different objects, and that selectivity is largely maintained across changes in object retinal position, even when images are simply presented to a fixating animal (*14*, *15*).

We tested a strong, "online" form of the temporal contiguity hypothesis—two monkeys visually explored an altered visual world (Fig. 1A, "Exposure phase"), and we paused every ~15 min to test each IT neuron for any change in position tolerance produced by that altered experience (Fig. 1A, "Test phase"). We concentrated on each neuron's responses to two objects that elicited strong (object "P", preferred) and moderate (object "N", nonpreferred) responses, and we tested the position tolerance of that object selectivity by briefly presenting each object at 3° above, below, or at the center of gaze (*16*) (fig. S1). All neuronal data reported in this study were obtained in these test phases: animal tasks unrelated to the test stimuli; no attentional cueing; and completely randomized, brief presentations of test stimuli (*16*). We alternated between these two phases (test phase ~5 min; exposure phase ~15 min) until neuronal isolation was lost.

To create the altered visual world ("Exposure phase" in Fig. 1A), each monkey freely viewed the video monitor on which isolated objects appeared intermittently, and its only task was to freely look at each object. This exposure "task" is a natural, automatic primate behavior in that it requires no training. However, by means of real-time eye-tracking (*17*), the images that played out on the monkey's retina during exploration of this world were under precise experimental control (*16*). The objects were placed on the video monitor so as to (initially) cast their image at one of two possible retinal positions (+3° or −3°). One of these retinal positions was pre-chosen for targeted alteration in visual experience (the "swap" position; counterbalanced across neurons) (Fig. 1B) (*16*); the other position acted as a control (the "non-swap" position). The monkey quickly saccaded to each object (mean: 108 ms after object appearance), which rapidly brought the object image to the center of its retina (mean saccade duration 23 ms). When the object had appeared at the non-swap position, its identity remained stable as the monkey saccaded to it, typical of real-world visual experience ("Normal exposure", Fig. 1A) (*16*). However, when the object had appeared at the swap position, it was always replaced by the other object (e.g., P→N) as the monkey saccaded to it (Fig. 1A, "Swap exposure"). This experience manipulation took advantage of the fact that primates are effectively blind during the brief time it takes to complete a saccade (*18*). It consistently made the image of one object at a peripheral retinal position (swap position) temporally contiguous with the retinal image of the other object at the center of the retina (Fig. 1).

We recorded from 101 IT neurons while the monkeys were exposed to this altered visual world (isolation held for at least two test phases; $n = 50$ in monkey 1; 51 in monkey 2). For each neuron, we measured its object selectivity at each position as the difference in response to the two objects (P − N; all key effects were also found with a contrast index of selectivity) (fig. S6). We found that, at the swap position, IT neurons (on average) decreased their initial object selectivity

McGovern Institute for Brain Research and Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge, MA 02139, USA.

*To whom correspondence should be addressed. E-mail: dicarlo@mit.edu

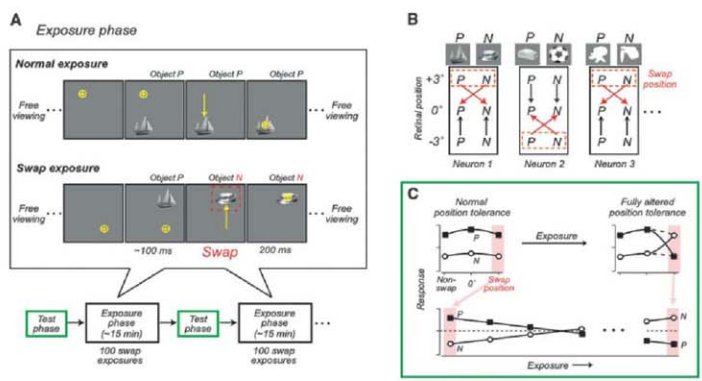


Fig. 1. Experimental design and predictions. **(A)** IT responses were tested in "Test phase" (green boxes, see text), which alternated with "Exposure phase." Each exposure phase consisted of 100 normal exposures (50 P→P, 50 N→N) and 100 swap exposures (50 P→N, 50 N→P). Stimulus size was 1.5° (*16*). **(B)** Each box shows the exposure-phase design for a single neuron. Arrows show the saccade-induced temporal contiguity of retinal images (arrowheads point to the retinal images occurring later in time, i.e., at the end of the saccade). The swap position was strictly alternated (neuron-by-neuron) so that it was counterbalanced across neurons. **(C)** Prediction for responses collected in the test phase: If the visual system builds tolerance using temporal contiguity (here driven by saccades), the swap exposure should cause incorrect grouping of two different object images (here P and N). Thus, the predicted effect is a decrease in object selectivity at the swap position that increases with increasing exposure (in the limit, reversing object preference), and little or no change in object selectivity at the non-swap position.

for P over N, and this change in object selectivity grew monotonically stronger with increasing numbers of swap exposure trials (Fig. 2, A and C). However, the same IT neurons showed (Fig. 2A) no average change in their object selectivity at the equally eccentric control position (non-swap position), and little change in their object selectivity among two other (nonexposed) control objects (see below).

Because each IT neuron was tested for different amounts of exposure time, we first computed a net object selectivity change, $\Delta(P-N)$, in the IT population by using the first and last available test phase data for each neuron. The prediction was that $\Delta(P-N)$ should be negative (i.e., in the direction of object preference reversal), and greatest at the swap position (Fig. 1C). This prediction was borne out (Fig. 3A). The position specificity of the experience-induced changes in object selectivity was confirmed by two different statistical approaches: (i) a direct comparison of $\Delta(P-N)$ between the swap and non-swap position ($n = 101$; $P = 0.005$, one-tailed paired t test); (ii) a significant interaction between position and exposure—that is, object selectivity decreased at the swap position with increasing amounts of exposure [$P = 0.009$ by one-tailed bootstrap; $P = 0.007$ by one-tailed permutation test; tests were done on $(P-N)$].

The changes in object selectivity at the swap position were also largely shape-specific. For 88 of the 101 neurons, we monitored the neuron's selectivity among two control objects not shown to the animals during the exposure phase (chosen similar to the way in which the P and N objects were selected, fully interleaved testing in each test phase) (*16*). Across the IT population, control object selectivity at the swap position did not significantly change (Fig. 2A), and the swap object selectivity changed significantly more than the control object selectivity (Fig. 3B) ($n = 88$, $P = 0.009$, one-tailed paired t test of swap versus control objects at the swap position).

These changes in object selectivity were substantial (average change of ~5 spikes/s per 400 exposures at the swap position) (Figs. 2C and 3C) and were readily apparent and highly significant at the population level. In the face of well-known Poisson spiking variability (*19*, *20*), these effects were only weakly visible in most single IT neurons recorded for short durations, but were much more apparent over the maximal 1-hour exposure time that we could hold neurons in isolation (Fig. 2C, lower panels). To determine if the object selectivity changes continued to grow even larger with longer periods of exposure, we next recorded multi-unit activity (MUA) in one animal (monkey 2), which allowed us to record from a number of (nonisolated) neurons around the electrode tip (which all tend to have similar selectivity) (*21*, *22*) while the monkey was exposed to the altered visual world for the entire experiment (~2 hours) (*16*). The MUA data replicated the single-unit results—a change in object selectivity only at the swap position (Fig. 2C) ("position × exposure" interaction: $P = 0.03$, one-tailed bootstrap; $P = 0.014$, one-tailed permutation test; $n = 10$). Furthermore, the MUA object selectivity change at the swap position continued to increase as the animal received even more exposure to the altered visual world, followed a very similar time course in the rate of object selectivity change (~5 spikes/s per 400 exposures) (Fig. 3C), and even showed a slight reversal in object selectivity (N > P in Fig. 4D).

Our main results were similar in magnitude (Fig. 3, A and B) and statistically significant in each of the two monkeys (monkey 1: $P = 0.019$; monkey 2: $P = 0.0192$; one-tailed t test). Each monkey performed a different task during the test phase (*16*), suggesting that these neuronal changes are not task dependent.

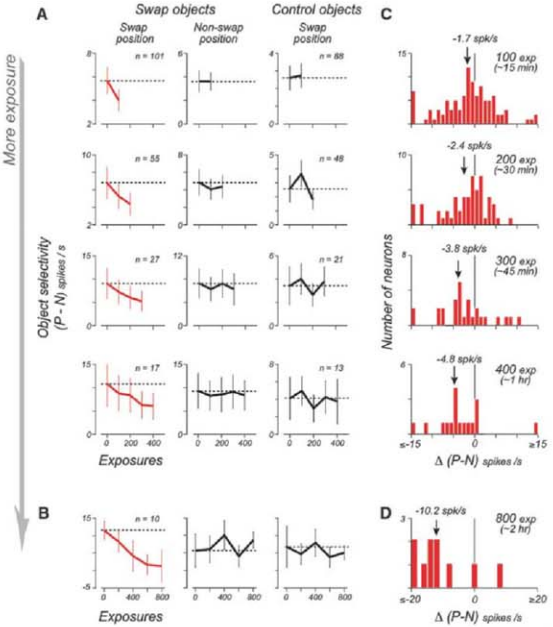


Fig. 2. Change in the population object selectivity. **(A)** Mean population object selectivity at the swap and (equally eccentric) non-swap position, and for control objects at the swap position. Each row of plots shows effect among all neurons held for at least the indicated amount of exposure (e.g., top row shows all neurons held for more than 100 swap exposures—including the neurons from the lower rows). The object selectivity for each neuron was the difference in its response to object P and N. To avoid any bias in this estimate, for each neuron we defined the labels "P" (preferred) and "N" by using a portion of the pre-exposure data (10 repetitions) to determine these labels, and the reminder to compute the displayed results in all analyses using these labels. Though there was, by chance, slightly greater initial selectivity at the swap position, this cannot explain the position specificity of the observed change in selectivity (table S2). **(B)** Mean population object selectivity of 10 multi-unit sites. Error bars (A and B) are SEMs. **(C)** Histograms of the object selectivity change at the swap position, $\Delta(P-N) = (P-N)_{\text{post-exposure}} - (P-N)_{\text{pre-exposure}}$. The arrows indicate the means of the distributions. The mean $\Delta(P-N)$ at the non-swap position was −0.01, −0.5, −0.9, and −0.9 spikes/s, respectively. The variability around that mean (i.e., distribution along the x axis) is commensurate with repeated measurements in the face of known Poisson spiking variability (fig. S11). **(D)** Object selectivity changes at the multi-unit sites. The mean $\Delta(P-N)$ at the non-swap position was 1.6 spikes/s.

Because we selected the objects P and N so that they both tended to drive the neuron (*16*), the population distribution of selectivity for P and N at each position was very broad [95% range: (−5.7 to 31.0 spikes/s) pooled across position; n = 101]. However, our main prediction assumes that the IT neurons were initially object-selective (i.e., the response to object P was greater than to object N). Consistent with this, neurons in our population with no initial object selectivity at the center of gaze showed little average change in object selectivity at the swap position with exposure (fig. S5). To test the learning effect in the most selective IT neurons, we selected the neurons with significant object selectivity [n = 52 of 101 neurons; two-way analysis of variance (2 objects × 3 positions), $P < 0.05$, significant main object effect or interaction]. Among this smaller number of object-selective neurons, the learning effect remained highly significant and still specific to the swap position ($P = 0.002$ by t test; $P = 0.009$ by bootstrap; $P = 0.004$ by permutation test).

To further characterize the response changes to individual objects, we closely examined the selective neurons held for at least 300 exposures (n = 28 of 52 neurons) and the multi-unit sites (n = 10). For each neuron and site, we used linear regression to measure any trend in response to each object as a function of exposure time (Fig. 4A). Changes in response to P and N at the swap position were apparent in a fraction of single neurons and sites (Fig. 4A), and statistically significant object selectivity change was encountered in 12 of 38 (32%) instances (Fig. 4C) (*16*). Across our neuronal population, the change in object selectivity at the swap position was due to both a decreased response to object P and an increased response to object N (approximately equal change) (Fig. 4B). These response changes were highly visible in the single-units and multi-units held for the longest exposure times (Fig. 4D).

These changes in the position profile of IT object selectivity (i.e., position tolerance) cannot be explained by changes in attention or by adaptation (fig. S10). First, a simple fatigue-adaptation model cannot explain the position specificity of the changes because, during the recording of each neuron, each object was experienced equally often at the swap and non-swap positions (also see additional control in table S2). Second, we measured these object selectivity changes with briefly presented, fully randomized stimuli while the monkeys performed tasks unrelated to the stimuli (*16*), which argues against an attentional account. Third, both of these explanations predict response decrease to all objects at the swap position, yet we found that the change in object selectivity at the swap position was due to an increase in response to object N (+2.3 spikes/s per 400 swap exposures) as well as a decrease in response to object P (−3.0 spikes/s per 400 swap exposures) (Fig. 4). Fourth, neither possibility can explain the shape specificity of the changes.

We term this effect "unsupervised temporal slowness learning" (UTL), because the selectivity changes depend on the temporal contiguity of object images on the retina and are consistent with the hypothesis that the natural stability (slowness) of object identity instructs the learning without external supervision (*6*–*11*). Our current data as well as previous human object perception experiments (*12*) cannot rule out the possibility that the brain's saccade-generation mechanisms or the associated attentional mechanisms (*23*, *24*) may also be needed. Indeed, eye-movement signals are present in the ventral stream (*25*, *26*). The relatively fast time scale and unsupervised nature of UTL may allow rapid advances in answering these questions, systematically characterizing the spatiotemporal sensory statistics that drive it, and understanding if and how it extends to other types of image tolerance (e.g., changes in object scale, pose) (*27*, *28*).

IT neurons "learn" to give similar responses to different visual shapes ("paired associates") when reward is used to explicitly teach monkeys to associate those shapes over long time scales [1 to 5 s between images; see, e.g., (*29*, *30*)], but sometimes without explicit instruction (*31*, *32*). A top-down explanation of the neuronal selectivity changes in our study is unlikely because animals performed tasks that were unrelated to the object images when the selectivity was probed, and the selectivity changes were present in the earliest part of the IT responses (~100 ms; fig S4). But UTL could be an instance of the same plasticity mechanisms that underlie "paired associate" learning; here, the "associations" are between object images at different retinal positions (which, in the real world, are typically images of the same object). However, UTL may be qualitatively different because (i) the learning is retinal position-specific; (ii) it operates over the much shorter time scales of natural visual exploration (~200 ms); and (iii) it is unsupervised in that, besides the visual world, no external "teacher" was used to direct the learning (e.g., no association-contingent reward was used, but we do not rule out the role of internal "teachers" such as efferent eye-movement signals). These distinctions are important because we naturally receive orders-of-magnitude more such experience (e.g., ~10^8 unsupervised temporal-contiguity saccadic "experiences" per year of life).

Our results show that targeted alteration of natural, unsupervised visual experience changes the position tolerance of IT neurons as predicted by the hypothesis that the brain uses a temporal contiguity learning strategy to build that tolerance in the first place. Several computational models show how such strategies can build tolerance (*6*–*11*), and such models can be implemented by means of Hebbian-like learning rules (*8*, *33*) that are consistent with spike-timing–dependent plasticity (*34*). One can imagine IT neurons using almost temporally coincident activity to learn which sets of its afferents correspond to features of the same object at different positions. The time course and task independence of UTL are consistent with synaptic plasticity (*35*, *36*), but our data do not constrain the locus of plasticity, and changes at multiple levels of the ventral visual stream are likely (*37*, *38*).

We do not yet know if UTL reflects mechanisms than are necessary for building tolerant representations. But these same experience manipulations change the position tolerance of human object perception—producing a tendency to, for example, perceive one object to be the

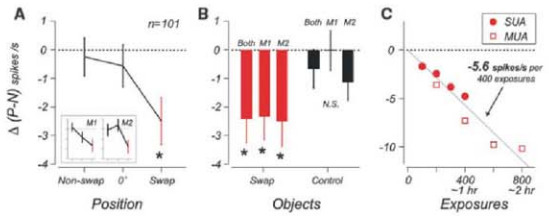


Fig. 3. Position specificity, object specificity, and time course. **(A)** Mean object selectivity change, Δ(P − N), at the swap, non-swap, and central (0°) retinal position. Δ(P − N) was computed as in Fig. 2C from each neuron's first and last available test phase (mean ~200 swap exposures). The insets show the same analysis performed separately for each monkey. **(B)** Mean object selectivity change for the (exposed) swap objects and (nonexposed) control objects at the swap position. Error bars (A and B) are SEMs. The swap object selectivity change at the swap position is statistically significant (*) in the pooled data as well as in individual animals ($P < 0.05$, one-tailed t test against 0). **(C)** Mean object selectivity change as a function of the number of swap exposures for all single-unit (n = 101) and multi-unit sites (n = 10). Each data point shows the average across all the neurons and sites held for a particular amount of time. Gray line is the best linear fit with a zero intercept; slope is mean effect size: −5.6 spikes/s per 400 exposures. The slope at the non-swap position based on the same analysis was 0.6 spikes/s (not shown).

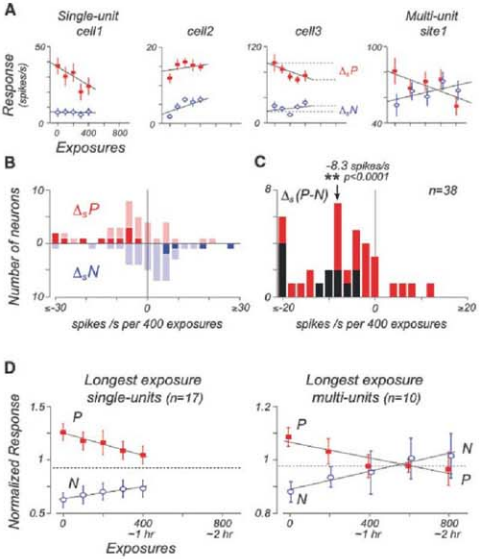


Fig. 4. Responses to objects P and N. **(A)** Response data to object P and N at the swap position for three example neurons and one multi-unit site as a function of exposure time. The solid line is standard linear regression. The slope of each line (Δ_S) provides a measure of the response change to object P and N for each neuron. Some neurons showed a response decrease to P, some showed a response enhancement to N, and others showed both (see examples). **(B)** Histograms of the slopes obtained for the object-selective neurons/sites tested for at least 300 exposures. The dark-colored bars indicate neurons with significant change by permutation test ($P < 0.05$) (*16*). **(C)** Histograms of the slopes from linear regression fits to object selectivity (P – N) as a function of exposure time; units are the same as in (B). Arrow indicates the mean of the distribution [the mean Δ_S(P – N) at the non-swap position was –1.7 spikes/s, $P = 0.38$]. The black bars indicate instances (32%; 12 of 38 neurons and sites) that showed a significant change in object selectivity by permutation test ($P < 0.05$). Results were very similar when we discarded neurons and sites with greater initial selectivity at the swap position (fig. S8). **(D)** Data from all the neurons and sites that were tested for the longest exposure time. The plot shows the mean normalized response to object P and N as a function of exposure time (compare to Fig. 1C; see fig. S3 for data at the non-swap position and for control objects). Error bars (A and D) are SEMs.

same identity as another object across a swap position (*12*). Moreover, given that the animals had a lifetime of visual experience to potentially build their IT position tolerance, the strength of UTL is substantial (~5 spikes/s change per hour)—just 1 hour of UTL is comparable to attentional effect sizes (*39*) and is more than double that observed in previous IT learning studies over much longer training intervals (*40*–*42*). We do not yet know how far we can extend this learning, but just 2 hours of (highly targeted) unsupervised experience begins to reverse the object preferences of IT neurons (Fig. 4D). This discovery reemphasizes the importance of plasticity in vision (*4*, *32*, *35*, *37*, *40*, *41*, *43*, *44*) by showing that it extends to a bedrock property of the adult ventral visual stream—position-tolerant object selectivity (*45*–*47*), and studies along the postnatal developmental time line are now needed.

References and Notes

1. S. Thorpe, D. Fize, C. Marlot, *Nature* **381**, 520 (1996).
2. M. C. Potter, *J. Exp. Psychol. (Hum. Learn.)* **2**, 509 (1976).
3. C. P. Hung, G. Kreiman, T. Poggio, J. J. DiCarlo, *Science* **310**, 863 (2005).
4. N. K. Logothetis, D. L. Sheinberg, *Annu. Rev. Neurosci.* **19**, 577 (1996).
5. R. Q. Quiroga, L. Reddy, G. Kreiman, C. Koch, I. Fried, *Nature* **435**, 1102 (2005).
6. L. Wiskott, T. J. Sejnowski, *Neural Comput.* **14**, 715 (2002).
7. P. Foldiak, *Neural Comput.* **3**, 194 (1991).
8. G. Wallis, E. T. Rolls, *Prog. Neurobiol.* **51**, 167 (1997).
9. R. Wyss, P. Konig, P. F. Verschure, *PLoS Biol.* **4**, e120 (2006).
10. T. Masquelier, S. J. Thorpe, *PLoS Comp. Biol.* **3**, e31 (2007).
11. T. Masquelier, T. Serre, S. J. Thorpe, T. Poggio, *CBCL Tech. Report #269*, Massachusetts Institute of Technology (2007).
12. D. D. Cox, P. Meier, N. Oertelt, J. J. DiCarlo, *Nat. Neurosci.* **8**, 1145 (2005).
13. G. Wallis, H. H. Bulthoff, *Proc. Natl. Acad. Sci. U.S.A.* **98**, 4800 (2001).
14. M. Ito, H. Tamura, I. Fujita, K. Tanaka, *J. Neurophysiol.* **73**, 218 (1995).
15. H. Op de Beeck, R. Vogels, *J. Comp. Neurol.* **426**, 505 (2000).
16. Materials and methods are available as supporting material on *Science* Online.
17. J. J. DiCarlo, J. H. R. Maunsell, *Nat. Neurosci.* **3**, 814 (2000).
18. J. Ross, M. C. Morrone, M. E. Goldberg, D. C. Burr, *Trends Neurosci.* **24**, 113 (2001).
19. D. J. Tolhurst, J. A. Movshon, A. F. Dean, *Vision Res.* **23**, 775 (1983).
20. M. N. Shadlen, W. T. Newsome, *J. Neurosci.* **18**, 3870 (1998).
21. K. Tanaka, *Cereb. Cortex* **13**, 90 (2003).
22. G. Kreiman *et al.*, *Neuron* **49**, 433 (2006).
23. T. Moore, M. Fallah, *Proc. Natl. Acad. Sci. U.S.A.* **98**, 1273 (2001).
24. E. Kowler, E. Anderson, B. Dosher, E. Blaser, *Vision Res.* **35**, 1897 (1995).
25. J. L. Ringo, S. Sobotka, M. D. Diltz, C. M. Bunce, *J. Neurophysiol.* **71**, 1285 (1994).
26. T. Moore, A. S. Tolias, P. H. Schiller, *Proc. Natl. Acad. Sci. U.S.A.* **95**, 8981 (1998).
27. S. Edelman, S. Duvdevani-Bar, *Neural Comput.* **9**, 701 (1997).
28. G. Wallis, H. Bulthoff, *Trends Cogn. Sci.* **3**, 22 (1999).
29. K. Sakai, Y. Miyashita, *Nature* **354**, 152 (1991).
30. A. Messinger, L. R. Squire, S. M. Zola, T. D. Albright, *Proc. Natl. Acad. Sci. U.S.A.* **98**, 12239 (2001).
31. Y. Miyashita, *Nature* **335**, 817 (1988).
32. C. A. Erickson, R. Desimone, *J. Neurosci.* **19**, 10404 (1999).
33. W. Gerstner, R. Kempter, J. L. van Hemmen, H. Wagner, *Nature* **383**, 76 (1996).
34. H. Sprekeler, C. Michaelis, L. Wiskott, *PLoS Comp. Biol.* **3**, e112 (2007).
35. C. D. Meliza, Y. Dan, *Neuron* **49**, 183 (2006).
36. H. Markram, J. Lubke, M. Frotscher, B. Sakmann, *Science* **275**, 213 (1997).
37. T. Yang, J. H. Maunsell, *J. Neurosci.* **24**, 1617 (2004).
38. Z. Kourtzi, J. J. DiCarlo, *Curr. Opin. Neurobiol.* **16**, 152 (2006).
39. J. H. R. Maunsell, E. P. Cook, *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **357**, 1063 (2002).
40. C. I. Baker, M. Behrmann, C. R. Olson, *Nat. Neurosci.* **5**, 1210 (2002).
41. E. Kobatake, G. Wang, K. Tanaka, *J. Neurophysiol.* **80**, 324 (1998).
42. N. Sigala, N. K. Logothetis, *Nature* **415**, 318 (2002).
43. E. T. Rolls, G. C. Baylis, M. E. Hasselmo, V. Nalwa, *Exp. Brain Res.* **76**, 153 (1989).
44. A. Seitz, T. Watanabe, *Trends Cogn. Sci.* **9**, 329 (2005).
45. M. Dill, M. Fahle, *Percept. Psychophys.* **60**, 65 (1998).
46. M. Dill, S. Edelman, *Perception* **30**, 707 (2001).
47. T. A. Nazir, J. K. O'Regan, *Spat. Vis.* **5**, 81 (1990).

48. We thank D. Cox, R. Desimone, N. Kanwisher, J. Maunsell, and N. Rust for helpful comments

and discussion, and J. Deutsch, B. Kennedy, M. Maloof, and R. Marini for technical support. This work was supported by the NIH (grant R01-EY014970) and The McKnight Endowment Fund for Neuroscience.

Supporting Online Material
www.sciencemag.org/cgi/content/full/321/5895/1502/DC1
Materials and Methods
SOM Text
Figs. S1 to 12
Tables S1 and S2
References and Notes

5 May 2008; accepted 15 August 2008
10.1126/science.1160028

Conformational Switch of Syntaxin-1 Controls Synaptic Vesicle Fusion

Stefan H. Gerber,[1*†] Jong-Cheol Rah,[2,3*‡] Sang-Won Min,[1*§] Xinran Liu,[1,4] Heidi de Wit,[5] Irina Dulubova,[6] Alexander C. Meyer,[3] Josep Rizo,[6,7] Marife Arancillo,[2] Robert E. Hammer,[6,7] Matthijs Verhage,[5] Christian Rosenmund,[2,3#] Thomas C. Südhof[1,4,8¶#]

During synaptic vesicle fusion, the soluble *N*-ethylmaleimide-sensitive factor–attachment protein receptor (SNARE) protein syntaxin-1 exhibits two conformations that both bind to Munc18-1: a "closed" conformation outside the SNARE complex and an "open" conformation in the SNARE complex. Although SNARE complexes containing open syntaxin-1 and Munc18-1 are essential for exocytosis, the function of closed syntaxin-1 is unknown. We generated knockin/knockout mice that expressed only open syntaxin-1B. Syntaxin-1B^{Open} mice were viable but succumbed to generalized seizures at 2 to 3 months of age. Binding of Munc18-1 to syntaxin-1 was impaired in syntaxin-1B^{Open} synapses, and the size of the readily releasable vesicle pool was decreased; however, the rate of synaptic vesicle fusion was dramatically enhanced. Thus, the closed conformation of syntaxin-1 gates the initiation of the synaptic vesicle fusion reaction, which is then mediated by SNARE-complex/Munc18-1 assemblies.

Intracellular membrane fusion reactions are carried out by interactions between SNARE [soluble *N*-ethylmaleimide–sensitive factor (NSF)–attachment protein receptor] and SM (Sec1-Munc18–like) proteins (*1*, *2*). In Ca^{2+}-triggered exocytosis in neurons and neuroendocrine cells, fusion is catalyzed by the formation of SNARE complexes from syntaxin-1, synaptosome-associated protein of 25 kDa (SNAP-25), and synaptobrevin/vesicle-associated membrane protein and the binding of the SM protein Munc18-1 to these SNARE complexes (*1*–*3*). Syntaxin-1 consists of two similar isoforms (syntaxin-1A and -1B) that are composed of an N-terminal α-helical domain (the H_{abc} domain) and a C-terminal SNARE motif and transmembrane region. Outside of the SNARE complex, syntaxin-1 assumes a "closed" conformation, in which the H_{abc} domain folds back onto the C-terminal SNARE motif (*4*, *5*). In the SNARE complex, by contrast, syntaxin-1 is "opened" (*6*). Munc18-1 interacts with syntaxin-1 alone in the closed conformation to form heterodimers (*3*, *4*) and additionally binds to SNARE complexes containing syntaxin-1 in the open conformation to form Munc18-1– SNARE complex assemblies (*7*, *8*), which are essential for exocytosis (*3*). The function of the closed conformation of syntaxin-1 and its binding to Munc18-1 remain unknown.

We used gene targeting to create mice that lack syntaxin-1A (syntaxin-1A^{KO}) and contain the LE mutation in syntaxin-1B, which renders it predominantly open (syntaxin-1B^{Open}) (fig. S1) (*9*). Studying littermate offspring from crosses of double-heterozygous syntaxin-1A^{KO} and -1B^{Open} mice, we found that homozygous syntaxin-1A^{KO} mice exhibited no decrease in survival (Fig. 1A) or other obvious phenotypes (figs. S2 and S3). The expendability of syntaxin-1A was unexpected in view of its high concentrations and proposed central functions (*10*–*14*) and indicated that syntaxin-1A may be functionally redundant with syntaxin-1B.

Homozygous mutant syntaxin-1B^{Open} mice were also viable but severely ataxic and developed lethal epileptic seizures after 2 weeks of age (Fig. 1A and fig. S3). The seizure phenotype of syntaxin-1B^{Open} mutant mice was recessive and independent of the syntaxin-1A^{KO}. Thus, syntaxin-1B was selectively essential, probably because it is more widely expressed than syntaxin-1A (*15*). In *Caenorhabditis elegans*, transgenic syntaxin-1^{Open} rescues unc-13 mutant worms from paralysis (*16*); however, crossing syntaxin-1B^{Open} mice with Munc13-1 knockout mice did not prevent Munc13-1 knockout–induced death (fig. S4).

The syntaxin-1A^{KO} mutation abolished syntaxin-1A expression (Fig. 1B), whereas the syntaxin-1B^{Open} mutation decreased syntaxin-1B levels (Fig. 1C). Both mutations produced a modest decrease in Munc18-1 levels but no major changes in other proteins (table S1). The syntaxin-1^{Open} mutation decreases formation of the Munc18-1–syntaxin-1 complex but not formation of SNARE complexes or Munc18-1–SNARE complex assemblies (fig. S5) (*3*, *8*). Consistent with this conclusion, less Munc18-1 was coimmunoprecipitated with syntaxin-1 in syntaxin-1B^{Open} mice, whereas other SNARE proteins coimmunoprecipitated normally (Munc18-1–SNARE complex assemblies are not stable during immunoprecipitations, and thus cannot be evaluated) (Fig. 1D and fig. S6).

Electron microscopy of cultured cortical neurons from littermate syntaxin-1B^{Open} and -1B^{WT} mice lacking syntaxin-1A revealed increased vesicle docking in syntaxin-1B^{Open} synapses (~25% increase) (Fig. 2, A to D). The size of the postsynaptic density also was increased (~20%) (Fig. 2E), whereas the density of docked vesicles per active zone length was unchanged (Fig. 2F). No other structural parameter measured differed between syntaxin-1B^{Open} and -1B^{WT} synapses; in particular, the number and intraterminal distribution of vesicles were unaltered (fig. S7). In chromaffin cells, however, the syntaxin-1B^{Open} mutation caused a large decrease in chromaffin vesicle docking, similar to that of the Munc18-1 knockout. Again, neither mutation altered the total number of chromaffin vesicles (Fig. 2, K and L). Synaptobrevin-2 knockout mice, analyzed in parallel as a negative control, did not change chromaffin vesicle docking but did increase the total number of chromaffin vesicles (Fig. 2L). Consistent with earlier findings (*17*–*20*), these results indicate that the Munc18-1–syntaxin-1 complex, but not the SNARE complex, functions in chromaffin vesicle docking. This function may not be apparent in vertebrate synapses because active zone proteins that are absent from chromaffin cells probably dock synaptic vesicles independent of their attachment to the Munc18-1–syntaxin-1 complex.

Measurements of spontaneous miniature excitatory postsynaptic currents (mEPSCs), excitatory postsynaptic currents (EPSCs) evoked by

[1]Department of Neuroscience, University of Texas Southwestern Medical Center, Dallas, TX 75390–9111, USA. [2]Department of Molecular and Human Genetics and Department of Neuroscience, Baylor College of Medicine, Houston, TX 77030, USA. [3]Department of Membrane Biophysics, Max-Planck-Institute for Biophysical Chemistry, 37077 Göttingen, Germany. [4]Department of Molecular Genetics, University of Texas Southwestern Medical Center, Dallas, TX 75390–9111, USA. [5]Department of Functional Genomics, Vrije Universiteit, 1081 Amsterdam, Netherlands. [6]Department of Biochemistry, University of Texas Southwestern Medical Center, Dallas, TX 75390–9111, USA. [7]Department of Pharmacology, University of Texas Southwestern Medical Center, Dallas, TX 75390–9111, USA. [8]Howard Hughes Medical Institute, University of Texas Southwestern Medical Center, Dallas, TX 75390–9111, USA.

*These authors contributed equally to this work.
†Present address: Abteilung Innere Medizin III, Universität Heidelberg, 69120 Heidelberg, Germany.
‡Present address: Developmental Synaptic Plasticity Section, National Institute of Neurological Disorders and Stroke, Bethesda, MD 20892, USA.
§Present address: University of California, San Francisco, Mission Bay Campus, San Francisco, CA 94158, USA.
¶Present address: Department of Molecular and Cellular Physiology and Neuroscience Institute, Stanford University, Palo Alto, CA 94304–5543, USA.
#To whom correspondence should be addressed. E-mail: rosenmun@bcm.tmc.edu (C.R.); tcs1@stanford.edu (T.C.S.)

isolated action potentials, and use-dependent synaptic depression during high-frequency stimulus trains in hippocampal neurons revealed no significant difference between syntaxin-$1A^{KO}$ and wild-type (WT) synapses (*12*). In syntaxin-$1B^{Open}$ synapses (on the syntaxin-$1A^{KO}$ background), however, the mEPSC frequency was increased ~40%, and use-dependent depression of EPSCs was massively enhanced, although evoked EPSCs unexpectedly exhibited normal amplitude and kinetics (Fig. 3 and fig. S8).

The increased depression in syntaxin-$1B^{Open}$ synapses indicates that they exhibit a decreased readily-releasable vesicle pool (RRP), an increased release probability, and/or a decreased rate of refilling of the RRP after it is emptied. To test these possibilities, we measured the RRP by applying 0.5 M sucrose (*9*, *21*). The RRP was unchanged in syntaxin-$1A^{KO}$ synapses but decreased ~35% in syntaxin-$1B^{Open}$ synapses (Fig. 4, A and B), which is consistent with decreased syntaxin-1 and Munc18-1 levels in syntaxin-$1B^{Open}$ mice (Fig. 1C). Determination of the RRP size in the same synapses in which we monitored mEPSCs and evoked EPSCs (Fig. 3, A to F) allowed us to calculate the spontaneous vesicular release rate (as the ratio of mEPSC frequency to the RRP) and the vesicular release probability P_{vr} (as the ratio of EPSC and RRP charges). Both were increased more than two-fold in syntaxin-$1B^{Open}$ synapses (Fig. 4, C and D), augmenting the percentage of the RRP released by a single action potential from ~10% in syntaxin-$1B^{WT}$ to ~20% in syntaxin-$1B^{Open}$ synapses. Measurements of the refilling rate of the RRP, however, detected an increase, not a decrease (fig. S9).

Thus, opening syntaxin-1 facilitates the fusion of synaptic vesicles on the background of a smaller RRP without changing the recruitment of vesicles into the RRP. Consistent with this conclusion, we found that the syntaxin-$1B^{Open}$ mutation accelerates sucrose-induced release (Fig. 4, E and F) and significantly boosts the relative amount and fractional release rate induced at lower sucrose concentrations (Fig. 4, G to I). Moreover, the syntaxin-$1B^{Open}$ mutation increases the apparent Ca^{2+} sensitivity of neurotransmitter release (fig. S10) and occludes the phorbol-ester–induced potentiation of release (fig. S11). Overall, these results establish that although the RRP is smaller in syntaxin-$1B^{Open}$ synapses, their resident RRP vesicles are more fusogenic.

The closed conformation of syntaxin-1 performs three functions upstream of the canonical role of syntaxin-1 as a SNARE protein in membrane fusion: (i) Closed syntaxin-1, but not the SNARE complex, mediates vesicle docking in chromaffin cells but not in synapses. The same differential phenotype is observed upon deletion of Munc18-1 (*17*, *18*), suggesting that the Munc18-1–syntaxin-1 complex docks chromaffin but not synaptic vesicles (fig. S12). (ii) The closed syntaxin-1 conformation stabilizes syntaxin-1 and Munc18-1, whereas opening syntaxin-1 decreases syntaxin-1 and Munc18-1 levels and thereby lowers the RRP size. (iii) Opening syntaxin-1 accelerates the rate of synaptic vesicle fusion, accounting for the fulminant epilepsy observed in synaxin-$1B^{Open}$ mutant mice.

Ca^{2+} and sucrose trigger fusion of primed synaptic vesicles. Primed vesicles are thought to be suspended in a metastable state in which SNARE complexes are assembled but the bilayers have not yet fused (*22*). We propose that primed vesicles are associated with a variable number of assembled SNARE complexes and that this number dictates the sucrose- and Ca^{2+}-sensitivity of a given vesicle (fig. S12). To account for the synaptic phenotype of syntaxin-$1B^{Open}$ mutant mice, we hypothesize that the syntaxin-$1B^{Open}$ mutation increases the average number of assembled SNARE complexes per vesicle and thereby enhances their Ca^{2+} and sucrose sensitivity. On the other hand, the destabilization of syntaxin-1 and Munc18-1 by the syntaxin-$1B^{Open}$ mutation (Fig. 1) decreases the total number of primed vesicles and thus the RRP, even though the primed vesicles are more fusogenic. The decrease in RRP is not due to the increased spontaneous release rate because its spontaneous fusion rate is still over 100-fold less than the

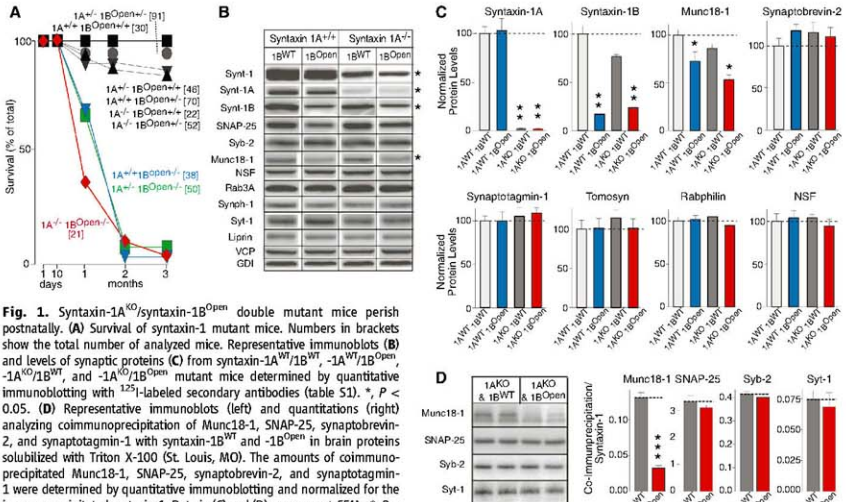


Fig. 1. Syntaxin-$1A^{KO}$/syntaxin-$1B^{Open}$ double mutant mice perish postnatally. **(A)** Survival of syntaxin-1 mutant mice. Numbers in brackets show the total number of analyzed mice. Representative immunoblots **(B)** and levels of synaptic proteins **(C)** from syntaxin-$1A^{WT}/1B^{WT}$, -$1A^{WT}/1B^{Open}$, -$1A^{KO}/1B^{WT}$, and -$1A^{KO}/1B^{Open}$ mutant mice determined by quantitative immunoblotting with ^{125}I-labeled secondary antibodies (table S1). *, $P < 0.05$. **(D)** Representative immunoblots (left) and quantitations (right) analyzing coimmunoprecipitation of Munc18-1, SNAP-25, synaptobrevin-2, and synaptotagmin-1 with syntaxin-$1B^{WT}$ and -$1B^{Open}$ in brain proteins solubilized with Triton X-100 (St. Louis, MO). The amounts of coimmunoprecipitated Munc18-1, SNAP-25, synaptobrevin-2, and synaptotagmin-1 were determined by quantitative immunoblotting and normalized for the immunoprecipitated syntaxin-1. Data in (C) and (D) are means ± SEMs; *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$ by Student's *t* test compared with wild type. Synt, syntaxin; Syb-2, synaptobrevin-2; Synph-1, synaptophysin-1; Syt-1, synaptotagmin-1; VCP, p97/vasolin-containing protein; GDI, guanosine diphosphate dissociation inhibitor.

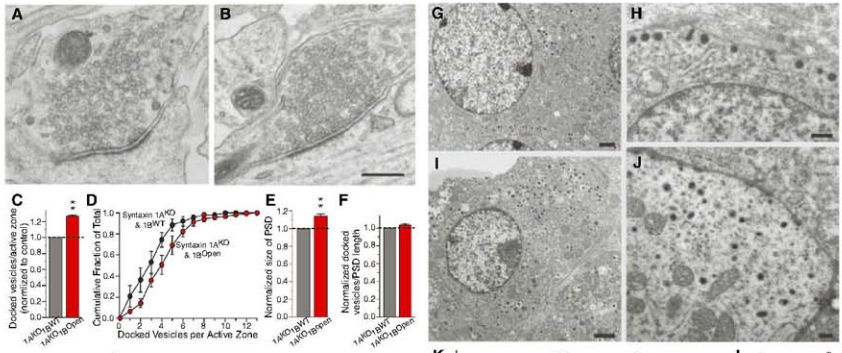


Fig. 2. Syntaxin-1B^{Open} impairs chromaffin but not synaptic vesicle docking. (**A** and **B**) Representative electron micrographs of neurons cultured from syntaxin-1A^{KO} mice containing wild-type syntaxin-1B^{WT} (**A**) or open syntaxin-1B^{Open} (**B**) (scale bar, 250 nm). (**C**) Number of docked vesicles per active zone (n = 3 experiments performed with 1B^{WT} = 49, 55, and 32 synapses and 1B^{Open} = 21, 38, and 49 synapses; normalized for wild-type values). (**D**) Plot of the cumulative distribution of docked vesicles per active zone (statistical significance with Kolmogorov-Simirnov test is $P < 0.01$; 1B^{WT} = 136 synapses and 1B^{Open} = 108 synapses). (**E** and **F**) Size of the postsynaptic density (PSD) (E) and number of docked vesicles/length of postsynaptic density (F) (both normalized for wild-type values). (**G** to **J**) Representative electron micrographs of chromaffin cells from control [(G) and (H)] and syntaxin-1B^{Open} [(I) and (J)] littermate mice at embryonic day E18 at two magnifications [(G) and (I), scale bar indicates 1 μm; (H) and (J), scale bar indicates 200 nm). (**K**) Distribution of the distance of secretory granules from the plasma membrane in chromaffin cells from syntaxin-1B^{WT}, -1B^{Open}, Munc18-1 knockout, and synaptobrevin-2 knockout mice (analyzed separately with wild-type controls and binned as indicated). (**L**) Total number of secretory granules per chromaffin cell in syntaxin-1B^{WT} or -1B^{Open} mice and in Munc18-1 and synaptobrevin-2 KO mice [(K) and (L), N = 3 animals, n = 60 chromaffin cells]. Data are means ± SEMs; **, $P < 0.01$ by Student's t test compared with wild type.

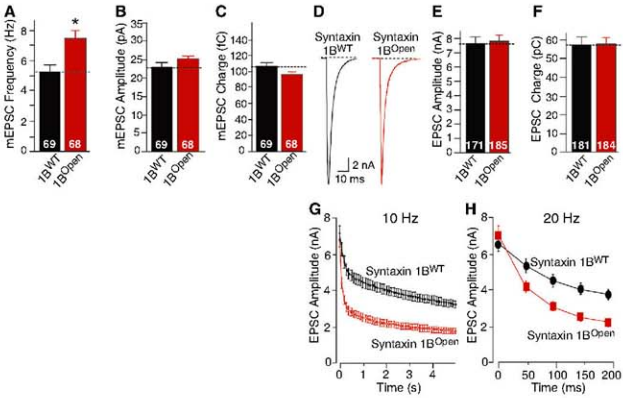


Fig. 3. Neurotransmitter release in syntaxin-1B^{Open} synapses. (**A** to **C**) Summary graphs of the frequency (A), amplitude (B), and charge (C) of spontaneous mEPSCs. (**D** to **F**) Representative traces (D), mean EPSC amplitudes (E), and charges (F) in synaptic responses induced by isolated action potentials. (**G** and **H**) EPSC amplitudes of evoked synaptic responses elicited by 10 Hz (G) and 20 Hz (H) stimulus trains. Data are means ± SEMs; *, $P <$ 0.05 by Student's t test compared with wild type. Numbers in bars show numbers of neurons analyzed.

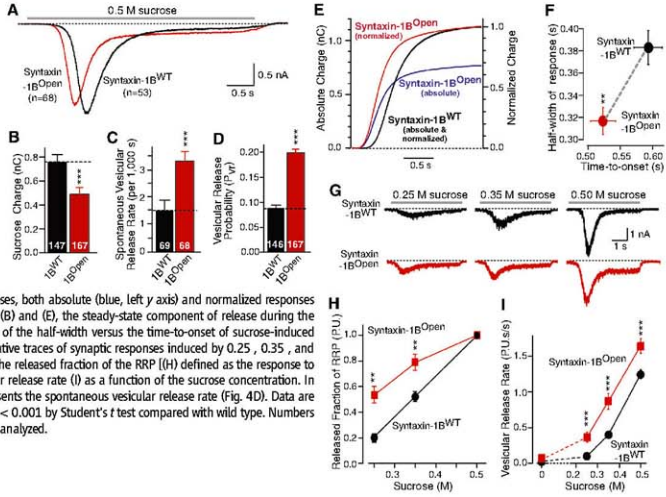


Fig. 4. Increased fusogenicity of synaptic vesicles in syntaxin-1B^{Open} synapses. **(A)** Average postsynaptic currents elicited by application of 0.5 M sucrose in syntaxin-1B^{WT} and -1B^{Open} synapses. **(B)** Mean RRP size determined as the transient charge integral induced by application of 0.5 M hypertonic solution. **(C)** Summary graph of the spontaneous vesicular release rate (minifrequency divided by the number of vesicles in the RRP). **(D)** Mean P_{vr}. **(E)** Time course of the average cumulative synaptic charge transfer during sucrose-induced release. For syntaxin-1B^{Open} synapses, both absolute (blue, left *y* axis) and normalized responses (red, right *y* axis) are depicted. In (B) and (E), the steady-state component of release during the responses was subtracted. **(F)** Plot of the half-width versus the time-to-onset of sucrose-induced synaptic responses. **(G)** Representative traces of synaptic responses induced by 0.25 , 0.35 , and 0.50 M sucrose. **(H** and **I)** Plot of the released fraction of the RRP [(H) defined as the response to 0.5 M sucrose] and of the vesicular release rate (I) as a function of the sucrose concentration. In (G), the 0 mM sucrose value represents the spontaneous vesicular release rate (Fig. 4D). Data are means ±SEMs; **, $P < 0.01$; ***, $P < 0.001$ by Student's *t* test compared with wild type. Numbers in bars show numbers of neurons analyzed.

vesicle repriming rate and because much higher spontaneous fusion rates in synaptotagmin-mutant mice do not decrease the RRP size (*23*). An alternative hypothesis would be that primed vesicles lack assembled SNARE complexes and Ca^{2+} or hypertonic sucrose trigger fusion of primed vesicles by inducing the opening of syntaxin-1 and assembly of SNARE complexes (fig. S12). The simplicity of this second model is attractive, but it cannot account for the speed of Ca^{2+}-triggered fusion or for its dependence on complexin, which binds to assembled SNARE complexes (*22*). Independent of which model is correct, our results demonstrate that syntaxin-1 performs multiple functions in exocytosis that go beyond its role as a SNARE protein to include the control of vesicle docking and the regulation of the vesicle fusion rate.

References and Notes

1. A. T. Brunger, *Q. Rev. Biophys.* **9**, 1 (2005).
2. R. Jahn, R. H. Scheller, *Nat. Rev. Mol. Cell Biol.* **7**, 631 (2006).
3. M. Khvotchev *et al.*, *J. Neurosci.* **27**, 12147 (2007).
4. I. Dulubova *et al.*, *EMBO J.* **18**, 4372 (1999).
5. K. M. Misura, R. H. Scheller, W. I. Weis, *Nature* **404**, 355 (2000).
6. R. B. Sutton, D. Fasshauer, R. Jahn, A. T. Brunger, *Nature* **395**, 347 (1998).
7. I. Dulubova *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **104**, 2697 (2007).
8. J. Shen, D. C. Tareste, F. Paumet, J. E. Rothman, T. J. Melia, *Cell* **128**, 183 (2007).
9. Materials and methods are available as supporting material on *Science* Online.
10. I. Bezprozvanny, R. H. Scheller, R. W. Tsien, *Nature* **378**, 623 (1995).
11. Z. H. Sheng, J. Rettig, T. Cook, W. A. Catterall, *Nature* **379**, 451 (1996).
12. A. P. Naren *et al.*, *Nature* **390**, 302 (1997).
13. S. L. Deken, M. L. Beckman, L. Boos, M. W. Quick, *Nat. Neurosci.* **3**, 998 (2000).
14. S. B. Condliffe, H. Zhang, R. A. Frizzell, *J. Biol. Chem.* **279**, 10085 (2004).
15. D. L. Foletti, R. Lin, M. A. Finley, R. H. Scheller, *J. Neurosci.* **20**, 4535 (2000).
16. J. E. Richmond, R. M. Weimer, E. M. Jorgensen, *Nature* **412**, 338 (2001).
17. M. Verhage *et al.*, *Science* **287**, 864 (2000).
18. T. Voets *et al.*, *Neuron* **31**, 581 (2001).
19. R. M. Weimer *et al.*, *Nat. Neurosci.* **6**, 1023 (2003).
20. H. de Wit, L. N. Cornelisse, R. F. Toonen, M. Verhage, *PLoS* **1**, e126 (2006).
21. C. Rosenmund, C. F. Stevens, *Neuron* **16**, 1197 (1996).
22. T. C. Südhof, *Annu. Rev. Neurosci.* **27**, 509 (2004).
23. Z. P. Pang, J. Sun, J. Rizo, A. Maximov, T. C. Südhof, *EMBO J.* **25**, 2039 (2006).

24. We thank N. Hamlin, I. Kornblum, A. Roth, I. Herfort, and E. Borowicz for technical assistance. This work was supported by grants from Deutsche Forschungsgemeinschaft (GE1042/1-1, GE1042/3-1, and GE1042/3-2 to S.H.G. and Ro1296/5-3 to C.R.), NIH (NS051262 to C.R. and N537200 to J.R.), and Netherlands Organization for Scientific Research (Gp0970-10-036 and 900-01-001 to M.V. and 916-36-043 to H.d.W.).

Supporting Online Material

www.sciencemag.org/cgi/content/full/1163174/DC1
Materials and Methods
Figs S1 to S13
Table S1 and S2
References

14 July 2008; accepted 5 August 2008
Published online 14 August 2008;
10.1126/science.1163174
Include this information when citing this paper.

Protein Immunoblot Imaging and Analysis

The FluorChem Q Imaging System provides the sensitivity of chemiluminescence and the quantitative power of fluorescence in one easy-to-use instrument. Designed with fast-lens technology, a peltier-cooled camera captures high-resolution images with a linear dynamic range that outperforms film, and with speeds 10 times faster than a laser scanner. Equipped with three integrated excitation sources for multicolor blots, the FluorChem Q is compatible with Cy dyes, Alexa dyes, Q dots, and chemiluminescent protein immunoblotting kits. This versatile system can also image fluorescent DNA and protein gels. The system comes with intuitive software that automatically stores all experimental imaging protocols as well as channel and filter settings.

Alpha Innotech
For information 510-483-962
www.alphainnotech.com

Fluorescence Stereomicroscopes

The Leica M205 FA and M165 FC stereomicroscopes are for demanding fluorescence applications in developmental, molecular, and cellular biology. The microscopes unite top performance zoom optics, resolution, and contrast to produce fluorescence images with outstanding richness of detail. The instruments feature patent-pending FusionOptics technology, which takes advantage of a neurological phenomenon: The left beam path produces great depth of field, while the right beam path provides a high-resolution image. The human brain then combines the best information from both channels, using it to compose an image whose resolution and depth of field have never been achieved in a stereomicroscope before. With its fully apochromatic optics, large zoom range, and resolution performance, the Leica M205 FA can show the viewer details that used to be invisible. The instruments also feature the TripleBeam principle: A patented third beam path reserved exclusively for fluorescence illumination delivers evenly illuminated, reflex-free fields of view at all zoom settings.

Leica
For information +49-(0)-6441-29-2550
www.leica-microsystems.com

3D Cell Proliferation Assay

Although more scientists are growing cells in a three-dimensional (3D) culture that mimics in vivo conditions and allows the cells to organize into complex structures, they have lacked a practical way to evaluate cell proliferation in these cultures. Four new 3D cell proliferation kits include a core kit and three others that contain basement membrane extract, laminin I, and collagen I. Designed for assessing the effects of compounds or genetic alteration on cell proliferation in 3D culture, the high throughput 96 stripwell formats allow one-step quantitative analysis and colorimetric detection. Protocols are provided for tumorgenicity and cytotoxicity.

Trevigen
For information 800-873-8443
www.trevigen.com

Palm-Size Thermocycler

The Palm PCR is a palm-size portable thermocycler with a wide dynamic range capable of delivering polymerase chain reaction (PCR) amplified product in as little as 15 minutes. Palm PCR is a battery-powered, stand-alone device that delivers efficient and fast amplification for nearly all kinds of DNA samples, including the human genome. Even a single copy of DNA can be amplified in as little as 25 minutes to an amount sufficient for typical agarose-gel detection. It can achieve uniformity of better than ±0.1°C and can run for more than four hours on a single charge of a lithium-ion battery.

Ahram Biosystems
For information 650-224-8575
www.ahrambio.com

Human Stem Cell Kits

New Human Stem Cell Nucleofactor Kits enable researchers to efficiently transfect various types of stem cells such as mesenchymal, hematopoietic, or neural stem cells without compromising their potential for differentiation. With these kits, for single reactions and up to 96-well transfections, transfection efficiencies of 60 percent to 80 percent with high viability can be achieved for stem cells. The kits' optimized transfection conditions allow the use of DNA, small interfering-RNA, or peptides.

Amaxa
For information +49-221-99199-464
www.amaxa.com

Immunodetection System

The Tropix Western-SuperStar Immunodetection System can detect proteins on protein immunoblots for up to 4.5 days after the substrate is added, compared with competitors' kits in which the signal is gone after 4 hours. The kits make use of light-emission kinetics that rivals horseradish peroxidase-based detection chemistry.

Applied Biosystems
For information 800-327-3002
www.appliedbiosystems.com/westernsuperstar

Pipetting Filters

Biohit Safe-Cone filters for the Biohit line of precision pipettors offer safe pipetting and prevent contamination of the pipettor and sample. The filters can prolong the life of the pipettor by protecting the internal mechanisms. Pipettor parts are also protected from vapors and liquids that could contaminate or damage internal components. The filters are available in standard form for general work and as Safe-Cone Plus format for radioactive, cell culture, and molecular biology applications.

Biohit
For information 800-922-0784
www.biohit.com

Electronically submit your new product description or product literature information! Go to **www.sciencemag.org/products/newproducts.dtl** for more information.

Newly offered instrumentation, apparatus, and laboratory materials of interest to researchers in all disciplines in academic, industrial, and governmental organizations are featured in this space. Emphasis is given to purpose, chief characteristics, and availabilty of products and materials. Endorsement by *Science* or AAAS of any products or materials mentioned is not implied. Additional information may be obtained from the manufacturer or supplier.

CREATIVE WAYS TO ENERGIZE YOUR CAREER

By day, Lynda Williams teaches physics at Santa Rosa Junior College in California. On weekends, audiences laugh at "Carbon Is a Girl's Best Friend" and other technology songs she writes and performs as "The Physics Chanteuse."

Like Williams, science faculty can transform their training and expertise into stimulating new professional activities, well beyond their job descriptions. While the responsibilities of a full-time position are always paramount, scientists, for example, might graft an innovative academic or research component onto existing duties, or earmark ample free time to explore another facet of their field. When an intriguing possibility arises, they can experiment with weaving that fresh pursuit into their professional spectrum. This article covers a few more "serious" ways in which faculty can get creative and perhaps increase the satisfaction level of the job. **By Carol Milano**

If a particular new dimension for a teaching portfolio intrigues the instructor, it's likely to appeal to students, also. While introducing an academic offering isn't quick or easy, it's definitely gratifying.

Create a Student Program

Current and prospective students kept asking the Regents Professor of Soil Science at Washington State University about courses in organic agriculture, which he has researched since 1985. Eager to encourage any farming interests, **John Reganold** developed a successful new "Organic Agriculture and Farming" course in 2001. Then, "one day I thought, why not just have a major?" says Reganold.

Colleagues at a mid 2002 faculty meeting encouraged him. For 18 months, he met repeatedly with chairpeople and professors in every Agricultural College department at WSU. "It's a science-based major—you're not just outside growing plants," stresses Reganold, who created specific soil science courses while weaving chemistry, biology, horticulture, entomology, economics, and other requirements into the curriculum. Each participating professor had to agree to shift 15 percent of an existing syllabus to organic agriculture. "We built on existing tradition, namely WSU's solid reputation in organic agriculture. I presented this as the mission of a land-grant institution."

Graduate assistant Kathi Peck sought farmland for the new major. Then a colleague, Cathy Perrillo, spent months on extensive national marketing surveys to provide justification. "Everything went smoothly, but it's time-consuming," Reganold recounts. "The package had to go through five [university] committees, like the Academic Senate. Then the State Higher Education Coordinating Board took several months to approve it."

Following approval in June 2006, America's first organic agriculture major quickly drew student and media attention. Advising most of the 16 current majors, Reganold relishes "seeing their excitement about what they're taking. Each is really interested in a particular path, like food or soil science, so we [follow] it. That's the biggest plus for me: to have any undergrad excited about their major is hard to do. Sometimes colleagues come and thank me. That's very rewarding."

Unlike Reganold, **William Kisaalita** didn't follow tradition. The University of Georgia professor of engineering has pioneered an unusual international program. By 1997, his fulfilling job and family life had generated "an urge to give something back," especially to help people like those he'd grown up with in Uganda. With economic development funding unlikely, and senior faculty urging a research focus while awaiting tenure, he had stifled his passion until professorship.

Hoping to involve undergraduates in international work, Kisaalita secured a small grant in 2000 to travel through Uganda, searching for possible study projects. Now, through his global component in the required Capstone Design Experience, small, carefully selected teams of fourth-year engineering students spend their summer or spring break in poor African villages. Their assignment: Solve a local problem, on site.

To persuade administrators, Kisaalita emphasized a 2001 pilot program's benefits. "If you can justify what students will learn, you can justify the project. Measurable value is the continued »

> "I get satisfaction from seeing the students being transformed, and from customers."

UPCOMING FEATURES

Focus on France— September 26
Top Employers — October 10
Focus on Ireland— October 31

CREDIT: (TOP) LYNDA WILLIAMS; (BOTTOM) THE ROCKEFELLER UNIVERSITY

Faculty Positions

"You have to choose where your expertise fits and can help"
–Gregory Knipp

best evidence." Now, the program integrates his research, teaching, and outreach responsibilities. "The blessing is, our solutions are so unique, they themselves are publishable." One successful invention was a solar-powered milk cooler that doubles production.

Villagers living on $2 a day recognize blessings, too. "We talk to them, they see us coming back, they develop a belief in what we're doing. They know we're here to stay, not just bringing students to see poor people, then write about them. I get satisfaction from seeing the students being transformed, and from the customers."

Reach for a Wider Audience

Beyond students and colleagues, who else could a professor's expertise benefit?

"We're not doing a good enough job getting science out to the public," laments **Louann Brizendine,** clinical professor at the University of California, San Francisco. To share her research findings with a wider audience, the neuropsychiatrist chose to write a mass-market book.

Fascinated by women's mental health since medical school, Brizendine began suspecting connections between female hormones and depression. Recruited by UCSF in 1988, she created a "Hormones in Psychiatry" course, inspired by her pregnancy and postpartum experiences. It evolved into her UCSF Women's Mood and Hormone Clinic, treating about 600 women annually. When patients taking a particular medication for depression experienced decreased libido, Brizendine measured testosterone and sexual function. Low levels of both generated her "passion to clarify biological aspects of women's health in hormonal issues."

She'd never imagined writing a book, until a fortuitous social encounter with an editor. *The Female Brain* was published five years later, in mid 2006. Brizendine worked on it rigorously, largely on weekends. Rewriting took a year, to present all the science more personally and less technically, through anecdotes and clear explanations. "What I thought the whole book would be is now Appendix I," she reflects.

Translated into 24 languages, *The Female Brain* has sold 300,000 copies in English alone. "The outcome way exceeded my expectations," admits Brizendine, gratified by letters she receives from high school students. "They'll say, 'I was never interested in science before. Now I want to know what courses I can take to be a neuropsychiatrist.' It's deeply meaningful." A sequel, *The Male Brain*, is due out in November 2008.

Other best-selling professors include Marion Nestle, Goddard Professor of Nutrition and Public Health at New York University, author of several well-reviewed books about food issues. Some scientists write mainstream newspaper, magazine, or e-articles. Lisa Sanders, has a popular monthly column, "Diagnosis," in *The New York Times Magazine*.

Apart from writing, academics can translate their knowledge into other modes, like practical talks for community or nonprofit groups. **Lynda Williams,** in her "Physics Chanteuse" guise, has presented her witty science songs at New York's Cornelia Street Café, the Swedish Arts and Science Festival, an American Physics Society conference, and the Inspiration of Astronomical Phenomena meeting in Palermo, Italy.

Develop a Small Business

Sometimes, pursuing a passion can evolve into an outside enterprise with a valuable product or service.

During a late '60s Peace Corps stint in Brazil, **Louis Kirchhoff** encountered Chagas disease, a major cause of Latin American morbidity and death, affecting 12 million to 14 million people. His early research soon brought a prestigious biomedical prize and four-year US National Institutes of Health fellowship. Now the professor of infectious disease, epidemiology and internal medicine at the University of Iowa - Carver College of Medicine, he has spent 75 percent of his time on research since 1985, funded by the Veterans Administration, American Heart Association, and other sources.

Often asymptomatic, Chagas is easily transmitted by blood transfusions. Seeking a seriodiagnostic tool to avoid infection, Kirchhoff and technician Keiko Otsu developed a simple, accurate test using recombinant DNA technology. Eventually negotiating all rights to his invention, he applied for US and several Latin American patents (see "Who's Rights?"). In 1998, with NIH small business funding, he set up Goldfinch Diagnostics, Inc., as an affiliate of the UI Technological Innovation Center.

In 2007, Goldfinch licensed his automated Chagas assay to Abbott Laboratories, which had originally contacted Kirchhoff years before, after learning of his research. The pharmaceutical firm is developing two products based on Kirchhoff's work.

What inspired him? "To have participated in developing something that will have an impact. If this works out, I'll donate blood soon, and it will be tested for Chagas—for which I provided the technical basis!" Kirchhoff adds, "I'm not Mother Theresa with a stethoscope—the idea of an effortless royalty stream every quarter is additional motivation."

Expert Advice

Another academic sideline is consulting. For **Gregory Knipp,** associate professor of industrial and physical pharmacy at Purdue University, separating consulting projects from academic responsibilities is delicate. Sought out because of his publications and reputation, Knipp was initially "enticed by trying to develop something new, plus some financial incentives." continued »

Featured Participants

Purdue University
www.purdue.edu

Rockefeller University
www.rockefeller.edu

Santa Rosa Junior College
www.santarosa.edu

University of California, San Francisco
www.ucsf.edu

University of Georgia
www.uga.edu

University of Iowa
www.uiowa.edu

Washington State University
www.wsu.edu

CREDIT: PURDUE NEWS SERVICE PHOTO/DAVID UMBERGER

"We're now a major player in NCI's long-term initiative to diagnose and cure cancer."
–David Soll

In pharmaceutical consulting, he tackles immediate problems—such as promising therapeutic agents with previously unnoticed barriers—and seeks ways to "salvage new compounds or delivery systems and bring them to market. It can really cut into weekends and nights, so I don't do it all the time. You have to choose where your expertise fits and can help." Though high fees can significantly supplement academic income, "I'd rather do something where I know I could be of benefit than something outside my league, just for a paycheck," Knipp reflects.

Some institutions allow professors one day a week for outside projects. They can have consulting checks sent to the school, for use as unrestricted funds (unlike a grant), for equipment, research, or travel. Consultants may help companies make critical decisions that could impact overall success potential, or recommend cutting an unrealistic project, says Knipp, convinced that his students benefit from the new perspectives he hones with each consulting project.

Expand the Scope of Your Work

Adapting expertise to an unfamiliar research area can be stimulating.

In 1986, with minimal expectations, **David Soll** assigned a corner in his lab to a small, faltering NIH program, the Developmental Studies Hybridoma Bank (DSHB). In 22 years Soll, the University of Iowa's Carver/Emil Witschi Professor of Biological Sciences, has transformed DSHB into a vast, nonprofit resource for animal cell researchers using hybridomas (hybrid cells bred by fusing an antibody-producing lymphocyte with a tumor cell). In 2006, for 65,000 customers worldwide, DSHB staff filled nearly 9,000 orders. "At commercial prices, our revenues would total $25 million, but we [charge] 10 percent to 15 percent of those fees," says Soll, proud that DSHB, operating six days a week, is completely self-funding.

DSHB broke even and started expanding after one year. "The hybridomas themselves became fascinating. As we began making and using them, I realized I was facilitating research, the Bank was paying for everything, and everyone in my lab was working for them. Suddenly I was becoming an expert on antibodies. It was intriguing—I loved it," Soll recounts.

Recently, DSHB expanded into microbial research, providing deeply discounted monoclonal antibodies and hybridomas. "The new branch will soon catch up to the older one, doing the same amount of good for society," Soll predicts. In 2007, NIH National Cancer Institute's new Proteomics Initiative selected DSHB as official bank and distributor for 20,000 hybridomas secreting antibodies against proteins encoded by one-quarter of the entire human genome: about 5,000 cancer-linked genes. "We're now a major player in NCI's long-term initiative to diagnose and cure cancer."

Like Soll, **Leslie Vosshall** built her career around tiny lab organisms, but expanded her focus differently. Using fruit flies, her neurogenetics and behavior laboratory at Rockefeller University researches the brain's processing of olfactory signals for food, danger, or potential mating partners. In 2003, she and her postdoctoral fellow, Andreas Keller, were eager to test a controversial new theory about scent's effects. Vosshall opted to undertake her first clinical trial. Fortunately, Rockefeller University Hospital encourages and assists basic scientists to research human subjects, explains Vosshall, the Chemers Family Associate Professor.

Recruiting was uncomplicated : "People are fascinated by smell, and our studies are not terribly invasive, requiring just a blood sample." After a three-hour screening, subjects perform smell tests, reporting reactions verbally or by computer during successive visits. Now on their fifth human trial, her team recently published a lengthy paper on the work.

Vosshall relishes distinctions between her research subjects. "You can communicate with humans! Fruit flies don't talk—we do things indirectly, assessing from behavior what we think happened. We do genetic experiments with fruit flies, but can't ask specific men and women to breed." Surprisingly, Vosshall finds human research more cost effective. "Fruit flies are donated but we have to feed them and pay people to take care of them. Longer term, it's more technical than having humans sniff odor vials."

Thanks to some foundation funding, Vosshall expects to continue researching both species, rather than choose either. "Our goal is to understand in both insects and humans how the brain perceives odor stimuli. So many scientists spend their whole career working on one model organism, it's very rewarding to have the opportunity to jump up the family tree."

So what steps to take to broaden horizons and invigorate a career? Be alert to possibilities that catch a particular interest; don't be afraid to consider adding a new dimension. Scientific faculty can be expert witnesses, join editorial boards, help plan professional conferences, establish a research consortium with other institutions—the possibilities are as boundless as ingenuity. The rewards, too, are infinite. As Soll of NIH confides, having assisted 30,000 scientists, "The excitement is because it's all intermingled with freedom, research, mission, and a feeling of accomplishment. The whole idea is to keep doing new things."

Carol Milano is an independent journalist in New York City, covering health care and science.

10.1126/science.opms.r0800059

Whose Rights?

What if the scientific work leads to a new device, procedure, or product? A salaried professor or researcher isn't completely independent.

Investigate the institution's policies about intellectual property. Start by checking with the department chair, or contact the Technology Innovation Center (or equivalent).

Louis Kirchhoff, a professor at University of Iowa's Carver College of Medicine, developed an assay test for a tropical disease. "When I recognized that I was generating intellectual property, I submitted a conflict disclosure because it belonged to the UI Research Foundation," he recounts.

Each situation is distinctive. If William Kisaalita's engineering students devise a solution for an African village's difficulties, "the University of Georgia has a policy on products developed in a course," he notes. But because his international on-site program has external funding, it's exempt.

For basic information on the legalities of claiming rights to one's own creations, visit the US Patent Office website (www.uspto.gov).

CREDIT: UNIVERSITY OF IOWA DEPARTMENT OF BIOLOGY